University of Ljubljana

Faculty of Electrical Engineering

Peter Rogelj

Non-Rigid Registration of Multi-Modality Images

Ph.D. Thesis

Supervisor: prof. Stanislav Kovačič, Ph.D.

Cosupervisor: prof. James C. Gee, Ph.D.

Ljubljana, 2003

Contents

Abstract **vii**

Povzetek **ix**

Znanstveno področje in problematika x

Deformabilna poravnava slik xii

Večmodalno merjenje lokalne podobnosti xiii

Modeli deformacij xiv

Izvirni prispevki k znanosti xvi

Točkovne mere podobnosti xvi

Simetrična poravnava xvii

Modeli deformacij xviii

Vrednotenje večmodalnih netogih poravnav xix

Sistem za poravnavo medicinskih slik xx

1. Introduction and Summary **1**

1.1 The definition of image registration 3

1.2 Related work on non-rigid registration 4

1.2.1 Feature space 4

1.2.2 Spatial deformation models 5

1.2.3 Multi-modality non-rigid Registration 7

1.3 Continuum mechanic non-rigid registration 8

1.4 Contributions of the thesis 10

2. Point Similarity Measures **12**

2.1 Introduction 12

2.1.1 Locality of similarity measures 13
2.2 The basic principle of point similarity measures 15
2.3 Multi-modality point similarity measures 19
2.3.1 Entropy based point similarity measure 19
2.3.2 MI based point similarity measure 21
2.3.3 Other point similarity measures 22
2.4 Point similarity measures and segmentation 23
2.4.1 Joint intensity distributions 23
2.4.2 Segmentation based point similarity measure 25
2.4.3 Point similarity measures without segmentation 29
2.5 Benefits of point similarity measures 30
2.5.1 Locality and spatial deformation models 31
2.5.2 Interpolation artifacts 32
2.6 Comparison of point similarity measures 37
2.6.1 Selection of point similarity measures 37
2.6.2 Comparison of similarity functions 39
2.6.3 Comparison of registration results 45
2.7 Conclusion 50

3. Symmetric Image Registration **51**
3.1 Introduction 51
3.2 Symmetric image registration 54
3.2.1 Implementation details 57
3.3 Results 59
3.3.1 Registration of simple objects 61
3.3.2 Recovering synthetic deformations 61
3.3.3 Registration of real interpatient data 63
3.3.4 Discussion 64
3.4 Conclusion 67

4. Spatial Deformation Models **68**
4.1 Convolution models 68
4.1.1 Elastic models 69

4.1.2 Viscous fluid models 71
4.1.3 Incremental models 72
4.1.4 Combined elastic-incremental model 72
4.1.5 Comparison 74
4.2 Filter kernels 76
4.2.1 Elastic kernel 76
4.2.2 Gaussian kernel 76
4.2.3 Comparison 79
4.3 Discussion and conclusion 82

5. Evaluation of Multi-Modality Non-Rigid Registration 84
5.1 Introduction 84
5.2 Non-rigid registration issues 86
5.3 Evaluation methods 87
5.3.1 Recovering synthetic deformations 88
5.3.2 Segmentation based evaluation 92
5.3.3 Evaluation based on registered images 93
5.4 Three step evaluation strategy 94
5.5 Conclusion 97

6. Image Registration System 98
6.1 Introduction 98
6.2 Implementation 98
6.2.1 Rigid registration 99
6.2.2 Non-Rigid registration 101
6.2.3 Multiresolution strategy 102
6.3 Results 104
6.3.1 Rigid registration 106
6.3.2 Non-rigid registration 114
6.4 Conclusion 117

7. Conclusion 122

References **125**

Biography **135**

Publications **137**

Zahvala / Acknowledgements **141**

Izjava **143**

Abstract

Registration of medical images is becoming an important tool for medical treatment and medical analysis. By finding spatial relations between two or more images it combines their information, which is useful for observing changes in anatomy and/or function during time, for comparing subjects and for merging information of multiple images. In this thesis we focus on multi-modality non-rigid registration, which is used for detecting complex spatial relations between images of different modality. We propose several novel approaches and improvements to the existing solutions.

The basic and most significant difference between our approach and other multi-modality non-rigid registration approaches is the application of point similarity measures, which is the first contribution of this thesis. Point similarity measures have been designed for high dimensional multi-modality registration, where the main problem is detection of localized image discrepancies. Point similarity measures enable direct assessment of the most localized image discrepancies by measuring similarity of individual points, although the images may be of different modalities. Such an extreme locality is obtained by separating the process of measuring similarity into two steps. In the first step information of the whole images is used to derive a point similarity function, which is an estimate of intensity dependence between the images. In the second step the point similarity function is used to measure similarity of individual points. The fundamental question concerning point similarity measurement is how to derive a suitable point similarity function, which defines a point similarity measure. We present several point similarity measures, where some of them are derived from popular global multi-modality similarity measures, and some of them are designed according to the requirements of multi-modality non-rigid registration.

Point similarity measures enable several improvements of the registration process. One of them is better estimation of external forces, which drive the registration. We propose an approach called symmetric image registration, which solves the problem of asymmetry of similarity measurement, which leads to registration inconsistency and reduces the quality of registration. The symmetric image registration approach tends to improve the registration by establishing a symmetric image interdependence. The symmetry is obtained by treating both of the images in the same manner. Both of the images may be modeled by spatial deformation models, although in our implementation one of the images is fixed. The most distinctive feature is interaction between the images. Images interact through

forces, in accordance with the Newton's third law of motion. Forces on one image reflect in opposing forces on the other image, which forms the basis for the symmetry. Consequently, the registration is driven by information obtained by measuring similarity in both registration directions, which improves registration consistency and registration correctness.

Point similarity measures do not use any spatial information. This makes estimation of external forces independent to the spatial deformation model. The functional independence of these two registration stages enables them to be designed independently to each other, although they are linked by an iterative registration process. The functional independence of the stages is important, because spatial regularization is now solely in the domain of spatial deformation model, which gains full control over the transformation properties and can better suit to the deformable properties of imaged tissues. We have analyzed different convolution based spatial deformation models, and propose a new model that combines elastic and incremental approaches. The proposed model reduces the systematic error of the elastic model and improves anatomical suitability of the incremental models.

Functional independence of the registration stages also contributes to evaluation of the overall registration system. We propose a three-step evaluation procedure, which improves the evaluation of the whole system by separately evaluating each of the two registration stages. This is especially useful in the case of multi-modality registration, where only a limited number of features, required for assessing the actual transformation, can be identified in both of the images.

All the proposed approaches are employed in our image registration system. It does not require any specific knowledge of the anatomy and can be used for any kind of 3D medical images. It was used for testing the proposed approaches, and for performing several registration task for the purpose of medical research. The system is fully functional, such that it is appropriate for demonstrating the capabilities of non-rigid, as well as rigid image registration. However, the system is not yet appropriate for clinical applications because it is not validated for specific registration tasks.

This thesis contributes to the field of medical image registration, which is a rapidly growing scientific area. A variety of clinical applications stimulate the development and improvement of registration approaches carried out by numerous researchers and research groups worldwide. The solutions presented in this thesis do not try to fit to any specific clinical task. Instead, they deal with general problems, which are common to all high-dimensional non-rigid registration tasks. As such they could be applied to specific clinically important applications.

Povzetek

Netoga poravnava večmodalnih medicinskih slik

Medicinske slikovne tehnike omogočajo zajem raznovrstnih podatkov, ki opisujejo zgradbo in funkcijo tkiv in organov. Bogata slikovna vsebina pomembno prispeva k odkrivanju in razumevanju bolezenskih stanj in sprememb. Pomembno vlogo pri tem ima poravnava slik, ki z določitvijo prostorske preslikave med slikami odpravi geometrijska neskladja in tako omogoča združevanje in izločanje medicinsko pomembne slikovne informacije.

Zaradi elastičnih lastnosti tkiv, različnih pogojev pri zajemanju slik, aktivnosti organov in patoloških sprememb so prostorske relacije med slikami lahko zelo kompleksne. Poravnavo takšnih slik omogočajo postopki netoge pravnave, ki zaznajo tudi prostorsko omejena lokalna neskladja in jih odpravijo z ustrezno deformacijo slik.

Poseben pomen ima poravnava slik različnih modalnosti, ki opisujejo različne lastnosti slikanih tkiv. Takšna večmodalna poravnava naprimer omogoča združevanje informacije o zgradbi organov, zajete z anatomskimi slikovnimi postopki kot so naprimer MRI, CT in rentgen, z informacijo o aktivnosti tkiv, zajeto s funcionalnimi slikovnimi tehnikami kot so naprimer fMRI, PET ali SPECT. Na ta način pridobljena informacija o aktivnosti anatomskih struktur je lahko ključnega pomena za medicinsko diagnostiko in načrtovanje zdravljenja.

Eden najzahtevnejših problemov poravnave je večmodalna netoga poravnava slik, ki združuje obe prej omenjeni lastnosti, netogost in večmodalnost. Vzrok za njeno dodatno zahtevnost je v tem, da običajni večmodalni postopki niso zmožni zaznavati lokalnih slikovnih neskladij, kar pa je pogoj za uspešno netogo poravnavo. Temu problemu smo se posvetili v tej doktorski disertaciji in razvili izvirne postopke, ki omogočajo natančno večmodalno netogo poravnavo slik ter izboljšajo lastnosti in uporabne vrednosti postopkov netoge poravnave slik.

Znanstveno področje in problematika

Tomografske slike so pomemben dejavnik sodobne medicinske diagnostike. V splošnem jih lahko razdelimo na anatomske in funkcionalne. Anatomske omogočajo vpogled v strukturo organov, medtem ko funcionalne opisujejo aktivnost tkiv. Razvoj medicinskih slikovnih tehnik je hiter, saj informacijska vsebina slik pomembno prispeva k odkrivanju in razumevanju bolezenskih stanj in sprememb. Iz slik želimo izluščiti čim več medicinsko pomembnih informacij, pri čemer pa je zelo pomembna ustrezna obdelava slik. Eden od pomembnih postopkov obdelave slik je poravnava slik. Poravnava slik omogoča združevanje informacijske vsebine dveh (ali več) slik ter tako pripomore k izločanju medicinsko pomembnih informacij, pomembnih za medicinsko diagnostiko ter načrtovanje zdravljenja.

Naloga poravnave slik je najti geometrijsko transformacijo, ki preslika prostor ene, poravnavane slike, v prostor druge, referenčne slike, tako da doseže optimalno prostorsko skladnost anatomskih struktur, ki jih sliki opisujeta. Na ta način torej dosežemo, da se iste anatomske strukture na obeh slikah nahajajo v istih legah, kar minimizira geometrijska neskladja med slikama. Tako poravnane slike omogočajo primerjanje slik zajetih v različnih časovnih obdobjih, pod različnimi pogoji, ali celo med različnimi pacienti, ter združevanje informacij pridobljenih z različnimi slikovnimi tehnikami. Poravnava slik je zato uporabljana za najrazličnejše medicinske aplikacije, kot naprimer izdelavo populacijskih anatomskih atlasov, anatomske in funkcionalne primerjave pacientov, ugotavljanje razvoja bolezni ter uspešnosti zdravljenja, načrtovanje radioterapij in operacij, medoperativno lokalizacijo anatomskih struktur in podobno.

Poravnavo slik lahko v grobem razdelimo na togo poravnavo in netogo poravnavo. Toga poravnava je namenjena iskanju prostorske odvisnosti med celotnima slikama na osnovi določanja ustreznega premika in rotacije ene od slik. Na ta način pogosto ni mogoče odpraviti vseh slikovnih neskladij, saj so le ta lahko tudi posledica elastičnosti tkiv, delovanja organov, različnih pogojev pri zajemanju slik ali celo patoloških sprememb, kar se odraža v deformaciji ene od slik. Odpravljanju takšnih slikovnih neskladij je namenjena netoga poravnava slik.

Poravnavi slik različnih modalnosti, ki so zajete z različnimi slikovnimi postopki in opisujejo različne lastnosti slikanih tkiv, je namenjena večmodalna poravnava slik. Ker ima vsaka modalnost slik drugačen medicinski pomen, je z večmodalno poravnavo slik pridobljeno dodatno znanje o medsebojni prostorski odvisnosti lastnosti, ki jih slike opisujejo. Kot primer naj omenimo večmodalno poravnavo CT in PET slik. CT slike dobro opisujejo anatomsko zgradbo telesa, medtem ko PET slike prikazujejo njegovo aktivnost. S poravnavo obeh modalnosti pridobimo znanje o aktivnosti anatomskih struktur. Večmodalna poravnava je težavnejša od enomodalne poravnave, saj so relacije med lastnostmi, ki jih slike predstavljajo, v splošnem nepoznane in zato so nepoznane tudi relacije med svet-

lostmi različnih slik. Še več, relacije med svetlostmi slik so v splošnem nestalne in se spreminjajo s pogoji (in nastavitvami) pri zajemanju slik. Večmodalna poravnava zato zahteva posebne postopke poravnave, ki se samodejno prilagodijo dejanskim svetlostnim odvisnostim med slikami.

Posebno težavna je večmodalna netoga poravnava slik, ki združuje obe prej omenjeni lastnosti, netogost in večmodalnost. Problem večmodalne netoge poravnave je v tem, da uveljavljeni večmodalni postopki niso primerni za odpravljanje kompleksnih lokalnih slikovnih neskladij, prav tako pa so uveljavljeni visokodimenzijski postopki netoge poravnave niso primerni za poravnavo slik različnih modalnosti. Vzrok za to je v težavnem merjenju lokalne poravnanosti slik različnih modalnosti. Postopki za večmodalno merjenje poravnanosti namreč niso zmožni delovati lokalno, saj zahtevajo razmeroma velika področja slik, kajti kompleksne in vnaprej neznane relacije med svetlostmi slik ocenjujejo statistično. Za rešitev omenjenega problema večmodalne netoge poravnave slik je bilo zazvitih vrsto postopkov. Razdelili smo jih v štiri skupine:

1. Poravnava s preslikavo modalnosti [32]. Kadar je svetlostna odvisnost slik poznana, oziroma jo je iz poravnavanih slik mogoče oceniti, je mogoče sklepati o pojavnosti slike v drugi modalnosti. Na tak način pridobljene slike enakih modalnosti je nato mogoče poravnati z običajnimi monomodalnimi postopki za netogo poravnavo. Pravilnost preslikave modalnosti je odvisna od kompleksnosti svetlostne odvisnosti slik. S svetlostno preslikavo je mogoče realizirati le surjektivne svetlostne odvisnosti, ki predpostavljajo, da se strukturi, ki ju je na osnovi svetlosti moč razločiti na prvi sliki, po svetlosti ločita tudi na drugi sliki. Realne svetlostne odvisnosti so le redkokdaj surjektivne, kar ima za posledico napake v preslikavi in posledično tudi poravnavi.

2. Poravnava z globalnim ocenjevanjem poravnanosti [73, 55, 70]. Izboljšanje poravnanosti lokalnih neskladij ima za posledico tudi izboljšanje globalne poravnanosti celotnih slik. Z merjenjem globalne poravnanosti slik, z običajnimi večmodalnimi postopki, je torej mogoče zaznati tudi lokalne spremembe poravnanosti slik, potrebne za netogo poravnavo. Področja, katerih poravnanost se ne spreminja, pri tem prispevajo informacijo, ki pripomore k boljši oceni svetlostne odvisnosti slik. Slaba stran tega postopka je zelo velika računska zahtevnost, saj vsaka od številnih potrebnih meritev lokalne poravnanosti zahteva analizo celotne slike, torej tudi področij, ki se ob dani lokalni transformaciji sploh ne spremenijo. Računska zahtevnost je odvisna tudi od števila prostorskih stopenj poravnave, ki se z omogočanjem bolj lokalnih deformacij zelo hitro veča. Zato so postopki na osnovi globalnih mer podobnosti zaradi časovnih zahtev omejeni na manj kompleksne deformacije.

3. Poravnava z bloki (ang. block matching techniques) [9, 11, 24, 34, 41, 46].

Ker večmodalni postopki merjenja poravnanosti slik niso primerni za poravnavo izrazito majhnih slikovnih področij, uporablja postopek poravnave z bloki nekoliko večja slikovna področja. Poravnavano sliko tako razdeli na tako velike bloke, ki jih je še mogoče neodvisno in dovolj zanesljivo poravnati z referenčno sliko. Pri tem se običajno uporablja toga poravnava. Rezultati poravnav blokov se pripišejo njihovim središčem, na osnovi katerih se z interpolacijo izračuna netoga transformacija celotne slike. Zmožnost deformiranja slike je omejena z velikostjo uporabljenih področij, kar onemogoča poravnavo izrazito lokalnih slikovnih neskladij, zaradi uporabe toge poravnave blokov pa se na mestih večjih deformacij lahko pojavi tudi sistematska napaka poravnave.

4. Deformabilna poravnava slik. Ta skupina postopkov predstavlja nadgradnjo enomodalnih netogih poravnav [4, 6, 14, 28], ki izhajajo iz fizikalnih lastnosti deformabilnih materialov. Deformabilna poravnava temelji na ocenjevanju lokalne poravnanosti slik, ki ovira razvoj tovrstnih večmodalnih poravnav. Ker so večmodalni postopki ocenjevanja poravnanosti slik v osnovi globalni, se deformabilna poravnava za slike različnih modalnosti uporablja le redko. V literaturi smo zasledili le dva postopka, ki bi ju eventualno lahko uvrstili v to skupino poravnav [34, 38], a tudi ta dva sta zaradi ocenjevanja poravnanosti razmeroma velikih področij slik omejena le na manj kompleksne deformacije. Kljub temu moramo poudariti, da je lokalnost večmodalnega ocenjevanja poravnanosti mogoče izboljšati z uporabo dodatnega znanja. Zadostna količina dodatnega znanja lahko omogoči ocenjevanje poljubno majhnih področij slik, kar lahko večmodalno poravnavo po svojih zmožnostih približa ali celo izenači z enomodalnimi postopki, ki omogočajo odpravo izrazito lokalnih slikovnih neskladij. Problematiki lokalnega ocenjevanja večmodalne poravnanosti slik in večmodalni deformabilni poravnavi slik smo se posvetili v tej doktorski disertaciji.

Deformabilna poravnava slik

Deformabilna poravnava slik je postopek netoge poravnave, ki se zgleduje po fizikalnih lastnostih materialov, ki se deformirajo pod vplivom zunanjih sil. Poravnavana slika je modelirana kot deformabilno telo in izpostavljena prostorsko porazdeljenim silam, ki spremenijo njeno konfiguracijo in jo na ta način poravnajo z referenčno sliko. Zunanje sile so določene tako, da delujejo v smeri povečanja medsebojne podobnosti slik. Slika se deformacijam upira z notranjimi silami, ki ustrezajo uporabljenemu prostorskemu modelu deformacij. Končni rezultat poravnave ustreza ravnovesnemu stanju, pri katerem so zunanje sile nasprotne notranjim. Možna je tudi implementacija na osnovi energijske kriterijske funkcije. V tem primeru ocenjeno odstopanje med slikama predstavlja zunanjo energijo, ki nadomešča zunanje sile, transformacija modela deformacij pa predstavlja no-

tranjo energijo. Ravnovesno stanje ustreza stanju minimalne skupne energije sistema. Opisani postopek je leta 1981 predlagal Broit v svoji doktorski disertaciji [7], prvi praktični sistem tega tipa pa sta leta 1989 predstavila Bajcsy in Kovačič [4]. Omenjeni sistem je bil osnova za nadaljnji razvoj deformabilnih postopkov poravnave slik in veliko število tovrstnih sistemov poravnave.

Deformabina poravnava slik je iterativni postopek in sestoji iz dveh sklopov. Prvi sklop je namenjen določitvi zunanjih sil, ki vodijo poravnavo v smeri povečanja meadsebojne podobnosti med referenčno sliko in poravnavano sliko, drugi sklop pa je model deformacij, ki zunanje slike preslika v ustrezno prostorsko transformacijo. V vsaki naslednji iteraciji se zunanje sile določi ponovno, glede na že doseženo poravnanost slik.

Zunanje sile so običajno določene kot gradient podobnosti slik glede na parametre transformacije, kar ustreza gradientnemu postopku optimizacije. Parametri transformacije pri deformabilni poravnavi ustrezajo premikom kontrolnih točk poravnavanie slike. Podobnost slik mora odražati pravilnost poravnanosti slik, zato je postopek merjenja podobnosti odvisen od relacij med slikama. Ker zunanje sile delujejo lokalno, morajo biti odvisne od lokalne poravnanosti slik, zato je tudi podobnost med slikama običajno merjena lokalno. Za enomodalno poravnavo slik je podobnost najpogosteje merjena na osnovi razlike med svetlostjo istoležnih točk slik ali pa na osnovi lokalne križne korelacije med slikama. Izbira mer podobnosti za večmodalno poravnavo slik je težavnejša, saj so večmodalne mere podobnosti v osnovi globalne in ne omogočajo neposrednega ocenjevanja lokalne poravnanosti slik.

Večmodalno merjenje lokalne podobnosti

Večmodalne mere podobnosti [39, 40, 83, 85, 84, 18, 78] omogočajo primerjanje slik ne glede na njihovo svetlostno odvisnost, ki pred poravnavo običajno ni poznana. Zato večmodalne mere podobnosti temeljijo na statističnih postopkih, s katerimi implicitno ocenijo svetlostno odvisnost med slikama. Statistični postopki zahtevajo veliko število vzorcev, v tem primeru veliko število slikovnih točk, kar omejuje uporabo večmodalnih mer podobnosti na razmeroma velika področja slik.

Za izboljšanje lokalnisti večmodalnih mer podobnosti je bilo predlaganih več postopkov, ki temeljijo na izboljšanju statistične ocene vezane svetlostne porazdelitve, ki je osnova za izračun večmodalnih podobnosti in je običajno določena z normiranjem vezanega histograma svetlosti slik. Najpreprosteje se lokalnost večmodalnih mer izboljša z zmanjšanjem svetlostne ločljivosti. Omenjeni postopek ni primeren za slike, ki vsebujejo svetlostno podobna, a pomensko različna področja, ki lahko zaradi zmanjšanja svetlosne ločljivosti postanejo nerazločljiva. Izboljšavo omenjenega pristopa predstavlja ocenjevanje vezane svetlostne poradelitve z uporabo Parzenove cenilke [58]. Tudi ta postopek

zmanjša svetlostno razločljivost, a brez uporabe nelinearnih postopkov. Računsko učinkovito se postopek realizira s filtriranjem vezanega histograma. Manj popularni so postopki izboljšave lokalnosti z uporabo mer podobnosti, ki temeljijo na enodimenzionalni statistiki. Takšna mera je naprimer entropija svetlostne porazdelitve vsote ali razlike slik [67, 9]. K izboljšanju lokalnosti večmodalnih mer podobnosti prispeva tudi uporaba a priori znanj. Kadar je vezana svetlostna porazdelitev poravnanih slik znana vnaprej, je primerna uporaba 'likelihood' mer podobnosti [44, 65, 66]. Postopek je zanesljiv le kadar predpostavljena svetlostna porazdelitev slik ne odstopa od dejanske, sicer pa lahko privede do napak poravnave. Da bi se tem izognili, so bili predlagani postopki, ki predpostavljeno svetlostno porazdelitev slik kombinirajo z ocenjeno dejansko porazdelitvijo, tako da se za oceno podobnosti uporabi njuna utežena vsota [47]. Namesto a priori vezane porazdelitve se lahko uporabi tudi vezana porazdelitev svetlosti celotnih poravnavanih slik [52]. Možnosti za izboljšavo večmodalnih mer podobnosti se kažejo tudi v uporabi drugačnega dodatnega znanja, naprimer dodatnega znanja o segmentaciji slik [87], znanja pridobljenega iz predhodno poravnanih slikovnih baz in podobno.

Kljub vsem predlaganim izboljšavam je zanesljivo merjenje večmodalnih podobnosti ostalo omejeno na razmeroma velika področja slik, ki ne zagotavljajo zadostne lokalnosti za pravilno poravavo kompleksnih slikovnih neskladij. Postopki poravnave zato lokalno poravnanost slik pogosto ocenjujejo z merjenjem podobnosti večjih področij, podvrženih lokalnim geometrijskim transformacijam [73, 55, 70]. Za te postopke je značilna visoka časovna zahtevnost, ki zmanjšuje njihovo uporabno vrednost.

Modeli deformacij

Model deformacij služi preslikavi zunanjih sil v anatomsko smiselno transformacijo. Modeli deformacij zato modelirajo lastnosti realnih deformabilnih materialov, najpogosteje elastičnost [4, 13, 33, 34] ali viskoznost [25, 14, 6].

Najznačilnejša fizikalna lastnost tkiv je elastičnost, zato je za modeliranje transformacij pri poravnavi slik najpogosteje uporabljen elastični model deformacij. Povezavo med zunanjimi silami in pripadajočo transformacijo elastičnega material podaja Navier-Stokesova parcialna diferencialna enačba, ki je tudi osnova za elastični model. Za reševanje diferencialne enačbe elastičnosti je bil najprej predlagan relaksacijski postopek [4], ki je računsko zelo zahteven. Leta 1993 je bil predstavljen postopek reševanja diferencialnih enačb elastičnosti z metodo končnih elementov [29, 26]. Ideja izhaja iz mehanike, kjer se končni elementi uporabljajo za modeliranje mehanskih obremenitev togih konstrukcij. Uporaba končnih elementov preslika diferencialno enačbo elastičnosti v enostavnejšo matrično enačbo, poleg tega pa omogoča, da se področja velikih deformacij modelira

natančneje od ostalih področij slik, kar zmanjša potrebno število kontrolnih točk in na ta način zmanjša tudi računsko zahtevnost.

Za elastični model deformacij je značilna sistematska napaka, saj so za ohranjanje deformiranega stanja potrebne od nič različne zunanje sile, ki lahko nastanejo le kot posledica nepopolne poravnave slik. Sistematska napaka elastičnosti se veča z velikostjo deformacij, kar onemogoča poravnavo velikih slikovnih neskladij. Za zmanjšanje sistematske napake elastičnosti je bil najprej predlagan inkrementalni pristop [4], kasneje pa je bil vpeljan viskozni model deformacij [25, 14]. Viskozni model ni podvržen sistematski napaki in omogoča poljubno velike deformacije. Zaradi velikih razlik med lastnostmi viskoznega modela deformacij in realnimi lastnostmi bioloških tkiv, pa takšen model ne zagotavlja pravilnosti poravnave. Omenjeni mehanizem namreč omogoča velike deformacije, ne glede na njihovo anatomsko pravilnost oziroma smiselnost.

Reševanje parcialnih diferencialnih enačb elastičnosti ali viskoznosti je računsko zelo zahtevno. Pomemben korak v smeri zmanjšanja računske zahtevnosti predstavlja uporaba konvolucijskih filtrov. Bro-Nielsen in Gramkow sta leta 1996 predstavila konvolucijski filter za reševanje parcialne diferencialne enačbe viskoznosti [6]. Pri tem je jedro konvolucijskega filtra enako impulznemu odzivu materiala. Enak filter je lahko uporabljen tudi za reševanje parcialne diferencialne enačbe elastičnosti. Tak pristop pomeni 1000 kratno pohitritev, saj je za izračun rešitve potreben en sam prelet podatkov. Računsko zahtevnost modelov deformacij je možno še dodatno zmanjšati z uporabo separabilnih konvolucijskih filtrov [31]. Najpogosteje je uporabljan Gaussov filter, ki je pogosto predstavljen kot aproksimacija fizikalnih modelov elastičnosti in viskoznosti. V tej skupini je potrebno izpostaviti 'demonov' algoritem [80, 81, 60], ki izhaja iz zakonov termodinamike, zanj pa je značilna ugodna računska zahtevnost.

Anatomske strukture običajno sestojijo iz več različnih tipov tkiv, ki imajo različne fizikalne lastnosti. To upoštevajo tako imenovani biomehanski modeli deformacij [33, 23]. Tovrstni modeli v praksi še niso deležni večje pozornosti, saj zahtevajo segmentacijo poravnavanih slik, hkrati pa je težavno tudi določanje optimalnih parametrov modela. Omenimo naj še verjetnostni pristop (angl. probabilistic matching), ki dopušča uporabo dodatnega znanja o možnih deformacijah, s katerim opiše anatomsko variabilnost [56, 27, 3, 28, 86]. Takšno dodatno znanje lahko zmanjša verjetnost napačne poravnave, a zaradi velike kompleksnosti in variabilnosti človeške anatomije je uporaba postopka omejena z možnostmi izdelave verjetnostnih modelov. Tovrstni modeli so v glavnem omejeni na normalne anatomije, brez patologij.

V posebnih primerih lahko poravnava deluje tudi brez eksplicitno določenega modela deformacij [52]. Na možne deformacije slik namreč vpliva tudi izbira velikosti področij, uporabljenih v postopku ocenjevanja lokalne poravnanosti slik. Pri uporabi večjih slikovnih področij je namreč občutljivost na detajlne slikovne

razlike manjša, zato le teh ni mogoče primerno zaznati in odpraviti. Ocenjena deformacija tako izhaja le iz večjih, nižje frekvenčnih slikovnih razlik.

Izvirni prispevki k znanosti

V doktorski disertaciji smo raziskovali, razvili in vrednotili postopke za večmodalno netogo poravnavo slik. Izvirni prispevki k znanosti so sledeči:

- Razvoj nove družine mer podobnosti, imenovane točkovne mere podobnosti, ki omogoča večmodalno merjenje podobnosti poljubno velikih področij slik, vključno s področji velikosti posameznih točk.
- Razvoj postopka simetrične poravnave slik, ki izboljša konsistenco in pravilnost netogih poravnav.
- Analiza modelov deformacij in razvoj kombiniranega prostorskega modela deformacij, ki združuje lastnosti elastičnih in inkrementalnih modelov z namenom zmanjšanja napake poravave.
- Analiza postopkov vrednotenja netoge poravnave slik in razvoj novega postopka vrednotenja, ki omogoča boljše zaznavanje napak poravnave.
- Razvoj sistema za poravnavo slik, ki omogoča togo in netogo ter enomodalno in večmodalno poravnavo slik.

Izvirni prispevki k znanosti so podrobneje pojasnjeni v nadaljevanju.

Točkovne mere podobnosti

Poglavje 2

Večmodalna netoga poravnava je namenjena odpravljanju slikovnih neskladij različnih prostorskih razsežnosti, vključno z izrazito lokalnimi. Za detektiranje in eventualno odpravo prostorsko manjših neskladij je potrebno ocenjevanje lokalne skladnosti slik, ki lahko temelji na merjenju globalne podobnosti, to je podobnosti celotnih slik, ali na merjenju lokalnih podobnosti, to je podobnosti majhnih področij slik. Postopek na osnovi merjenja globalne podobnosti je zaradi velike računske in časovne zahtevnosti neprimeren za visoko dimenzionalne poravnave, in tako omejen le na odpravljanje enostavnejših slikovnih neskladij. Problem merjenja lokalnih podobnosti pa je v tem, da običajni in uveljavljeni postopki merjenja večmodalne podobnosti za to niso primerni, saj zaradi statističnega delovanja zahtevajo razmeroma velika pordočja slik. V principu uporaba manjših področij omogoča zaznavo prostorsko manjših slikovnih neskladij in tako omogoča

natančnejšo poravnavo slik. Da bi omogočili poravnavo vseh slikovnih neskladij ne glede na njihovo prostorsko razsežnost, smo razvili novo vrsto mer podobnosti, ki omogoča večmodalno merjenje podobnosti poljubno velikih področij slik, vključno z najmanjšimi, ki obsegajo le posamezno točko slike. Zaradi točkovnega ozadja te mere imenujemo točkovne mere podobnosti.

Točkovne mere podobnosti smo razvili iz uveljavljenih globalnih večmodalnih mer podobnosti, kot sta entropija in medsebojna informacija. Večmodalne mere morajo za oceno poravnanosti dveh slik implicitno oceniti tudi svetlostno odvisnost med slikama, ki je lahko zelo kompleksna in je običajno vnaprej nepoznana. Določitev svetlostne odvisnosti temelji na statistični oceni vezane porazdelitve svetlosti slik, za kar je potrebno veliko število vzorcev slike in zato razmeroma velika področja slik. Da bi omogočili merjenje podobnosti poljubno majhnih področij slik, smo proces merjenja večmodalne podobnosti razdelili v dva koraka: določitev svetlostne odvisnosti in dejansko merjenje podobnosti. Določitev svetlostne odvisnosti med poravnavanima slikama temelji na oceni vezane porazdelitve svetlosti celotnih slik. Možnih je več postopkov, ki ustrezajo različnim točkovnim meram podobnosti. Rezultat, izračunana svetlostna odvisnost, podaja vrednost podobnosti za vsak možen svetlostni par referenčne in poravnavane slike. V drugem koraku, je svetlosna odvisnost uporabljena za merjenje podobnosti posameznih točk slik. Podobnost dveh točk, točke na referenčni sliki in točke na poranavani sliki, je enaka podobnosti, ki jo svetlostna odvisnost določa za točkama pripadajoč svetlosni par. Podobnost večjih področij slik je enaka povprečju podobnosti točk (običajno vokslov) na danem področju.

Svetlostna odvisnost slik je lahko izračunana le enkrat in velja za vse točkovne pare poravnavanih slik. Razvili smo več različnih točkovnih mer podobnosti, ki se razlikujejo glede na način določitve svetlostne odvisnosti. Nekatere točkovne mere smo izpeljli iz uveljavljenih globalnih večmodalnih mer podobnosti, druge pa smo razvili na osnovi zahtev večmodalne netoge poravnave slik. Razvite točkovne mere podobnosti smo med seboj primerjali in jih preizkusili v sistemu večmodalne netoge poravnave slik.

Simetrična poravnava

Poglavje 3

Poravnava slik temelji na predpostavki, da večja podobnost med slikama pomeni boljšo poravnanost slik. Vendar pa mere podobnosti ne morejo zaznati transformacij, ki ne spremenijo izgleda poravnavane slike. V primeru netoge poravnave to predstavlja velik problem, saj onemogoča zanesljivo ocenjevanje poravnanosti informacijsko manj bogatih področij slik, naprimer svetlostno homogenih področij slik. Zunanje sile se zato lahko pojavijo le na informacijsko bogatih področjih, to je na področjih z od nič različnim gradentom svetlosti, naprimer na prehodih

med različnimi tipi tkiv. Vzrok tega je, da s poravnavo slik v različnih smereh, to je poravnavo prve slike na drugo sliko in poravnave druge slike na prvo sliko, dobimo različno poravnanost slik, saj se zunaje sile, ko sliki še nista poravnani, pojavijo na različnih mestih slik. Ta nesimetričnost kaže na slabo konsistenco in pravilnost poravnave. Za rešitev omenjenega problema smo predlagali postopek simetrične poravnave.

Simetrična poravnava je postopek, ki obe sliki obravnava na enak način. Sliki si delita skupni svetovni koordinatni sistem in se seveda lahko prekrivata. Vsaka od slik je lahko modelirana s svojim modelom deformacij, ki pa je lahko za vsako od slik drugačen. V skladu s tem se vsaka od slik lahko giblje in deformira. Najpomembnejša razlika glede na običajne postopke poravnave pa je interakcija med slikama. Sliki ena na drugo delujeta z zunanjimi silami, ki so po tretjem Newtonovem zakonu simetrične. Če ena od slik na drugo deluje z neko silo, potem druga slika deluje na prvo s po velikosti enako a nasprotno silo. Zunanje sile se tako pojavijo kot posledica zaznane neporavnanosti prve slike na drugo sliko, kot tudi druge slike na prvo sliko. Rezultat delovanja takšnih simetričnih sil je transformacija slik, ki celotni sistem privede v ravnovesno stanje prostorsko poravnanih slik. Takšna poravnava je zaradi simetričnosti postopka bolj konsistentna ter zaradi večje količine uporabljene informacije tudi bolj pravilna od poravnave po klasičnih, nesimetričnih postopkih.

Simetrično poravnavo slik smo učinkovito realizirali z uporabo točkovnih mer podobnosti. Pri tem smo referenčno sliko modelirali kot fiksno, poravnavano sliko pa z ustreznim deformabilnim modelom. Rezultate postopka simetrične poravnave smo primerjali z rezultati dveh običajnih nesimetričnih postopkov, pri čemer smo uporabili enostavne testne slike ter sintetične in realne medicinske slike. Rezultati dokazujejo prednosti simetrične poravnave.

Modeli deformacij

Poglavje 4

Prostorski model deformacij določa deformabilne lastnosti poravnave. Pričakovano je, da najbolj pravilno poravnavo slik dosežemo z uporabo fizikalnih modelov deformacij, ki se zgledujejo po deformabilnih lastnostih poravnavanih tkiv. Najznačilnejša lastnost tkiv je elastičnost, vendar pa je za elastične modele deformacij značilna sistematska napaka poravnave, ki onemogoča natančno poravnavo in je še posebno izrazita v primeru velikih deformacij. Kot rešitev tega problema sta se v praksi uvejavila dva druga modela deformacij, inkrementalni in viskozni. Oba omogočata na videz dobro poravnavo, brez sistematske napake, značilne za elastični model. Njune lastnosti pa se razlikujejo od lastnosti tkiv, zato v poravnavo vnašata napako, predvsem na informacijsko manj bogatih področjih slik. Da bi zmanjšali napako poravnave

smo razvili nov kombiniran model deformacij, ki združuje lastnosti elastičnih in inkrementalnih modelov. Elastični del pripomore k lineranejši odvisnosti med silami in doseženo deformacijo, inkrementalni pa k zmanjšanju sistematske napake elastičnosti.

Za dosego realističnih deformacij slik so potrebni ne le realistični modeli deformacij, temveč tudi realistične zunaje sile. Zunanje sile so v primeru netoge poravnave določene tako, da delujejo v smeri izboljšanja podobnosti slik, zato delujejo le na informacijsko bogata področja slik, kar običajno ustreza prehodom med različnimi tipi tkiv. Temu ustrazno je potrebno prilagoditi uporabljen model deformacij. Problem smo ilustrirali z uporabo modela deformacij na osnovi konvolucijskega filtriranja. Pokazali smo, da konvolucijsko jedro, ki ustreza dejanskim lastnostim tkiv, za poravnavo slik ni nujno najprimernejše. To opravičuje uporabo enostavnejših in računsko manj zahtevnih Gaussovih modelov. Modeli deformacij bi torej morali upoštevati ne le lastnosti tkiv, temveč tudi pričakovano porazdelitev sil, ki delujejo na poravnavane anatomske strukture.

Vrednotenje večmodalnih netogih poravnav

Poglavje 5

Rezultat poravnave slik je transformacija poravnavane slike, ki je potrebna za dosego prostorske skladnosti slikam pripadajočih anatomskih struktur. Ker pravilna transformacija realnih slik nikoli ni točno znana, je vrednotenje netogih poravnav v praksi zelo težavno, saj neposredna primerjava dosežene transformacije s pravilno transformacijo ni možna. Za vrednotenje netogih poravnav se zato uporabljajo indirektni postopki, ki pa imajo pomanjkljivosti in ne omogočajo zaznavanja vseh napak poravnave. Pri načrtovanju poravnav in njihovem vrednotenju se je potrebno zavedati pomanjkljivosti postopkov vrednotenja, saj boljši rezultat vrednotenja ne pomeni nujno tudi boljše poravnave. V tej disertaciji smo opozorili na slabosti posameznih postopkov vrednotenja ter predlagali nov način vrednotenja, primeren za poravnave na osnovi točkovnih mer podobnosti.

V primeru uporabe točkovnih mer podobnosti sestoji poravnava iz dveh funkcijsko ločenih sklopov: sklop za določanje zunanjih sil, katerega naloga je maksimiranje podobnosti slik, in prostorski model deformacij, ki zagotavlja anatomsko smiselnost poravnave. Zaradi funkcijske ločenosti je vsak sklop mogoče vrednotiti ločeno. Temu ustrezno predlagamo vrednotenje v treh korakih. V prvem koraku vrednotimo sklop določanja zunanjih sil, v drugem koraku vrednotimo model deformacij in v tretjem celoten sistem poravnave. Takšen način vrednotenja prinaša pomembno prednost predvsem v primeru večmodalne poravnave slik, saj je zaradi majhnega števila značilnic, ki jih lahko določimo na slikah različnih modalnosti, ocena realnih transformacij v tem primeru še posebej težavna. Vrednotenje posameznih sklopov je namreč manj zahtevno kot vred-

notenje celotne poravnave. Sklopa za določanje zunanjih sil namreč ni potrebno vrednotiti na osnovi realnih transformacij, temveč so te lahko generirane sintetično in tako točno znane. Prav tako vrednotenje modela deformacij ne zahteva slik različnih modalnosti, pač pa so lahko uporabljene slike iste modalnosti, na osnovi katerih se transformacija lahko oceni natančneje. Prva dva koraka vrednotenja tako omogočata boljše detekcijo napak, izbiro postopkov in nastavitev sistema kot to omogoča vrednotenje celotnega sistema. Vrednotenje celotnega sistema, kot tretji korak vrednotenja, je še vedno potrebno in zahteva večmodalne slike in realne deformacije, a skupna kvaliteta vrednotenja je boljša in omogoča zaznavanje več vrst napak poravnave.

Sistem za poravnavo medicinskih slik

Poglavje 6

Večmodalna netoga poravnava slik je namenjena uporabi v klinični medicini kot tudi za medicinske raziskave. Gre za razmeroma nov postopek, ki predvsem v klinični medicini še ni uveljavljen. Razlogov za to je več, med drugim tudi težavno vrednotenje, ki ni možno brez visoko usposobljenih strokovnjakov s področja medicine. Da bi širše predstavili zmožnosti netoge poravnave, omogočili vrednotenje predlaganih postopkov in olajšali nadaljne raziskave, smo razvili splošen sistem za poravnavo slik.

Sistem, ki smo ga razvili omogoča togo in netogo ter enomodalno in večmodalno poravnavo slik. Poravnavane so lahko slike najrazličnejših modalnosti in poljubnih delov telesa, iste ali različnih oseb. Sistem temelji na postopkih predlaganih v tej doktorski disertaciji. Uporabljene so točkovne mere podobnosti, za netogo poravnavo pa tudi postopek simetrične poravnave in kombiniran model deformacij. Sistem ne uporablja nobenega specifičnega znanja, kar prispeva k njegovi splšnosti. Razvit sistem smo uporabili za testirane predlaganih postopkov ter za vrsto poravnav namenjenih medicinskim raziskavam. Sistem omogoča nadaljne raziskave na področju obdelave medicinskih slik in je primeren za demonstracijo zmožnosti, ki jih poravnava nudi na področju medicinske diagnostike in načrtovanja zdravljenja. Sistem ni primeren za uporabo v klinični medicini, saj ni ustrezno vrednoten.

1. Introduction and Summary

Tomographic images play an important role in modern medical diagnosis. They enable observation and analysis of body structure and function. The information, which they provide, can be essential for detection and understanding of diseases and pathological differences. By registering multiple images, i.e. putting them into a spatial correspondence such that location of imaged anatomical structures is identical in both images, the information of the images is combined and useful extra information is obtained. This enables observation of changes in anatomy and/or function during time, comparison between multiple subjects and integration of information provided by different modality images. Registration can be used for various medical applications, e.g. development of anatomical atlases, comparison of subjects, medical examination, planning and evaluation of medical treatment, intraoperative localization, etc.

Different properties of imaged object/subject can be measured using different image acquisition techniques. Some examples of images acquired using different imaging techniques, i.e. images of different modality, are shown in Figure 1.1. The relation between measured properties is usually not known and, therefore, the relation between intensities of different images is not known either. This makes registration of such *multi-modality* images more complex than registration of *mono-modality* images, obtained using the same acquisition method. However, when multi-modality images are registered, valuable additional information can be obtained. For example, multi-modality registration of CT images, which show the anatomy, and PET images, which represent activity, provides the information about activity of anatomical structures.

Differences between the images can also appear due to the changes of geometry of anatomical structures. The geometrical differences in the anatomy can appear due to different position of the patient, organ activities or due to pathology. Furthermore, additional geometrical differences between the images may appear due to different imaging conditions. Such images can be correctly registered using *non-rigid registration* methods, which not only move and rotate the images, like in the case of *rigid registration*, but also deform them in such a way as tissues are expected to deform in reality. This, for example, enables detection and analysis of localized image differences as well as comparison and propagation of medical knowledge between different anatomies.

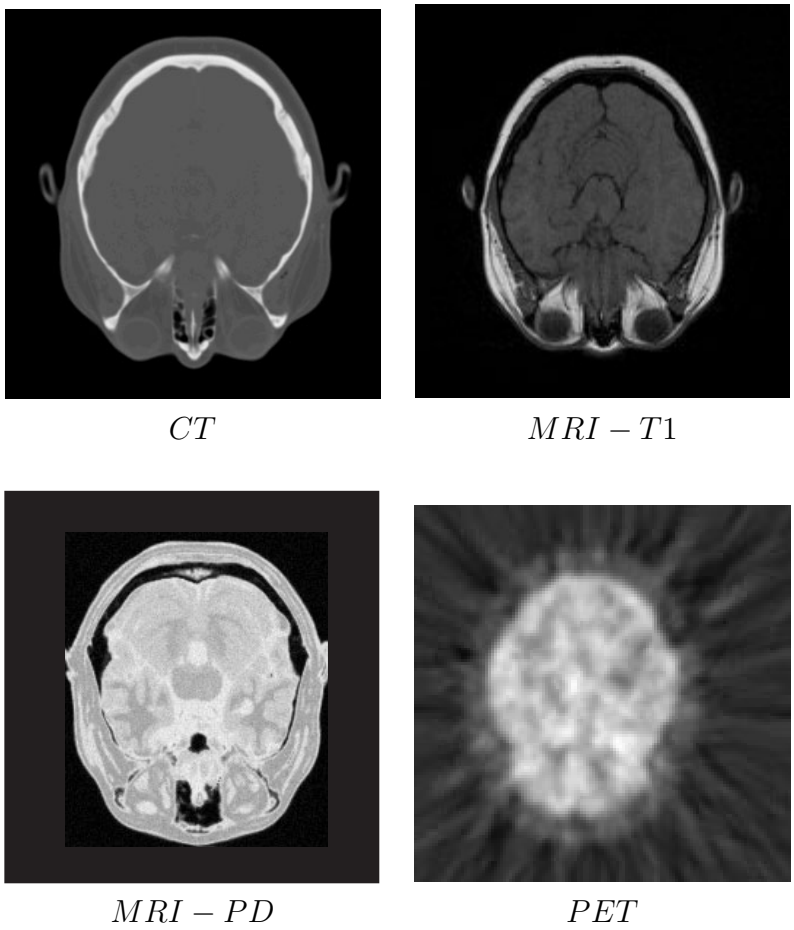


Figure 1.1: *Images of different modalities, acquired using different imaging techniques. Because they represent different physical properties of tissues, they look very different although they belong to the same anatomy. The relations between intensities of different modality images is not known, which makes multi-modality registration more complex than mono-modality registration.*

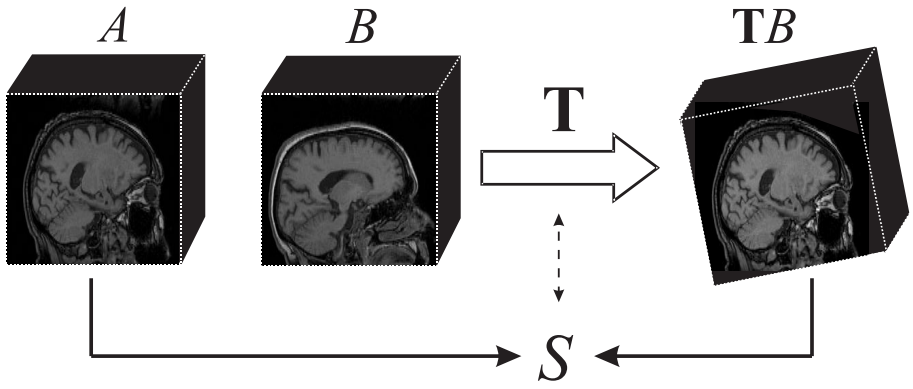


Figure 1.2: *Image registration is a process of finding transformation* $\mathbf{T}$*, which puts two imaged anatomies in correspondence. Search for such transformation is normally based on assumption that more correct transformation reflects in higher similarity* S *between target image* A *and transformed source image* $\mathbf{T}B$*.*

One of the most difficult registration tasks is *multi-modality non-rigid registration*. Complex and unknown multi-modality relations between intensities of involved images makes estimation of local image correspondence difficult. Consequently, precise registration of images with complex local image discrepancies is difficult as well. In this doctoral dissertation we deal with the problems of multi-modality non-rigid registration and propose new methods and improvements, which aid to the quality of registration results.

1.1 The definition of image registration

Image registration is a process, which involves two images, A and B, and tends to find a geometrical transformation $\mathbf{T}$ that maps one image to the other one, such that each anatomical point in one image gets moved to the same world coordinates as the corresponding anatomical point in the other image.

The correctness of the transformation usually cannot be measured directly, but it is commonly assumed that more correct registration results in more similar images. Consequently, in practice, the registration is a process, which searches for such transformation that maximizes some criterion function in which image similarity S plays an essential role, with the assumption that the obtained transformation is the one, which correctly matches the imaged anatomies. The situation is illustrated in Figure 1.2.

Unfortunately, the assumption that higher image similarity corresponds to more correct registration is not always valid. This is the most evident in the case of non-rigid registration. Here, the increase of similarity may also be related

to physically meaningless transformation. To prevent that, non-rigid registration must be spatially regularized, usually by spatial deformation model, which prevents such incorrect transformations.

1.2 Related work on non-rigid registration

In this section we review and classify the related work on non-rigid image registration and pay special attention to the work on multi-modality non-rigid registration. For more general review and classification of registration approaches see the papers of Brown [8], Lester and Arridge [43], Maintz [53], Pluim et al. [64], Thompson and Toga [82] and Zitova and Flusser [90].

Non-rigid image registration techniques in general differ according to the type of feature space used for estimating the image correspondence and according to the type of spatial deformation model used for regularizing the image transformation.

1.2.1 Feature space

According to the type of feature space used for estimating the image correspondence we classify non-rigid image registration techniques into *geometrically based approaches* and *intensity based approaches*.

Geometrically based approaches estimate the deformation field from localized geometric features. Such features include point landmarks [72, 71, 45], curves [49, 59] and surfaces [22, 21, 79], all of these can be localized manually or automatically. The resulting deformation is obtained by pairing these features in both images and interpolating the surrounding deformation field using splines, radial basis functions, diffusion or physical spatial deformation models. The quality of registration highly depends on quality of feature localization, which is a difficult task. Automatic feature extraction is difficult due to complex geometry and texture of anatomical features, while manual feature extraction is very time consuming, requires highly qualified medical experts, and still, it is subjective. This difficulties limit the use of geometrically based approaches.

Intensity based approaches employ the information of the whole images. In contrast to geometrically based approaches they do not require any feature extraction and they do not directly consider anatomical features. Image registration is performed as an optimization procedure that optimizes the transformation parameters in order to maximize the quality of image alignment. The quality of image alignment is measured using mathematical or statistical methods, which are called similarity measures. To achieve realistic registration results, without tearing or folding of the image, transformation is regularized using spatial defor-

mation models, which often model physical properties of deformable materials, e.g. elasticity [7, 4, 28], viscosity [14, 6] or simplified/fictious convolution based models [60, 73].

In general, intensity based approaches depend on larger number of image voxels than geometric approaches, which makes them more time consuming. However, they do not require feature extraction and enable registration of images where feature extraction is difficult or unfeasible. Due to this advantage and fast advancement of computers, intensity based registration approaches are more widely used, and are also the subject of this work.

1.2.2 Spatial deformation models

The rapid evolution of intensity based non-rigid registration techniques started in 1981, when Broit proposed modeling of images by physical deformable media [7]. The image, which is being registered, is treated like an elastic medium, which can be deformed by applying spatially distributed external forces. Following this principle, the registration is a process in which external forces, which are derived from the images, drive the transformation (deformation) process toward the final configuration, which in ideal case puts both of the images into register. One of the first systems of this kind was built in 1989 by Bajcsy and Kovačič [4]. Their *elastic deformation model* was based on Navier-Stokes equation, which links the external forces with consequent image deformation. They also proposed an *incremental* approach, which reduces the registration error caused by elasticity constraint. In order to avoid local minima and provide computational efficiency, they also used a coarse-to-fine strategy. This system has motivated research and development of other non-rigid registration systems, which have demonstrated several improvements.

In 1993 a *finite element* approach was proposed for solving partial differential equation of elasticity [29, 26]. Finite elements were adopted from mechanics, where they are used for computing mechanic loads of rigid objects. In the field of non-rigid medical image registration this approach aids to more realistic modeling of anatomies and enables estimation of actual body tensions. Furthermore, it enables *biomechanical modeling* of anatomy, such that different physical properties are used for different tissue types [33, 23]. However, biomechanical models depend on the registered anatomy and require image segmentation. Difficulties also arise in finding the optimal setting of mechanical properties. Consequently, biomechanical modeling has not attracted much attention yet.

Elastic models do not enable precise correction of large deformations. By increasing the required deformation the registration error increases as well. This systematic error can be reduced by using incremental deformations mentioned earlier. The drawback of incremental approach is in nonlinearity and reduced

consistency, which originates in non-linear relation between remaining image discrepancy and estimated external forces. One alternative to the elastic model is a *viscous fluid model* [25, 14], which enables large deformation without systematic error. This model differs from real tissue properties and thus it does not constrain the deformation realistically. Although it enables arbitrary large deformations, they are not necessarily anatomically correct and phisically meaningful.

Solving of partial differential equations (of elasticity or viscosity) is computationally complex and time consuming. An important step in reducing the computational requirements was made in 1996 by Bro-Nielsen and Gramkow, who developed a convolution filter to solve the system of partial differential equations in a single pass over the data [6]. The convolution kernel for elasticity equals the impulse response of the elastic media. This approach speeds up the computation by a factor of 1000 [82].

The computational complexity of spatial deformation models can be additionally reduced by using separable convolution filters [31]. The most often used kernel is Gaussian, which can be understood as an approximation of either elastic or viscous fluid models. Convolution filtering is also used in a popular 'demon' algorithm [80, 81, 60], which follows the principles of thermodynamics and is famous for its low computational cost.

In special cases the registration can be performed without explicit spatial deformation model [52]. The actual image deformation capability does not depend only on explicit deformation model, but also on the registration procedure, specifically on the size of regions used for estimating the quality of local image match. The larger the regions are, the lower is the sensitivity to detailed high frequency image differences, and thus the obtained deformations are smoother and more constrained.

We have to mention that external forces used in the registration process do not necessarily equal the forces that actually act on the deformed anatomy. Therefore, the results obtained using the previously mentioned physically inspired models may still not be absolutely correct. One important idea to improve the results is the use of *probabilistic approach*. In this case the non-rigid registration is considered as a problem of finding image transformation with the highest probability, using decision-theoretic approach, based on Bayesian modeling [56, 27, 3, 28, 86]. Probabilistic matching incorporates prior knowledge by using probability models, which specify the anatomic variability of admissible spatial and intensity transformations. Such additional knowledge is expected to improve the registration, but due to high anatomical variability it is difficult to build adequate probability models, which are usually limited to anatomies without pathology.

1.2.3 Multi-modality non-rigid Registration

Multi-modality non-rigid registration is considerably more complex than mono-modality non-rigid or multi-modality rigid registration. The problem arises because multi-modality similarity measures are due to their statistical nature all global, while non-rigid registration is used to deal with local image discrepancies. This prevents multi-modality non-rigid registration to directly follow the mono-modality non-rigid registration solutions, where similarity is measured locally. Thus, in the mono-modality case local image discrepancies can be directly assessed, mostly by measuring correlation or mean square intensity difference [8, 39, 66]. In contrast, the existing multi-modality similarity measures are in general all global and cannot be used locally. The most often used multi-modality similarity measures are mutual information, which was proposed in 1995 by Viola and Wells [83] and Collignon [18], and normalized mutual information, which was proposed in 1999 by Studholme et.al. [78]. Several authors have proposed novel approaches for multi-modality non-rigid registration in order to avoid the problem that appears due to the global nature of multi-modality similarity measures. We classify these approaches into four categories:

Registration approaches based on global similarity measures [73, 55, 70]. These approaches detect improvement of local image correspondence by observing global image similarity. They are suitable for simple smooth deformations, which can be modeled by low parameter deformation models. The drawback of these approaches is that each estimation of some local property requires recomputation of global image similarity i.e. similarity of the whole images. Consequently, by increasing the number of parameters this approaches become very time consuming, and therefore, they are not applicable for high dimensional problems.

Block matching techniques [11, 24, 34, 41, 46] divide one image into smaller regions, which are independently registered to the other image using multiple low-dimensional transformations. Transformations obtained for the individual image regions are then used to obtain smooth global transformation. In general, registration results improve by reducing the region size and increasing the number of regions. However, regions cannot be arbitrarily small, as reduction of region sizes decreases performance of multi-modality similarity measures used for estimating image correspondence. Therefore, these approaches are also not suitable for detecting complex local image discrepancies.

Approaches based on intensity transformation. One such approach was proposed by Guimond et al. [32]. The method is based on applying intensity transformation to one of the images, such that it matches the intensity properties of the other image. After that, images can be registered using high-dimensional mono-modality, instead of multi-modality, non-rigid registration algorithms. Such approaches can in theory reach the same accuracy as mono-modality approaches. However, they are limited to images with a functional intensity dependence. It

is required that intensity transformation is a surjective function, such that two different tissues that differ in intensity in one image do not have the same intensity representation on the other image. Real multi-modality images do often not comply to this requirement.

Continuum mechanics approaches [34, 38]. This category of approaches tends to follow the mono-modality continuum mechanics solutions, such as [4, 6, 14, 28]. For this purpose it is necessary to develop a method to obtain external forces for each voxel of the image that is being registered. Relying on the gradient descent optimization procedure, such external forces can be obtained as a gradient of image similarity. Here, it is possible to use the commonly used multi-modality similarity measures. Although they are in general global (cannot be applied to local image regions), they may still be locally sensitive (can detect local image differences). However, this approach is computationally complex and less appropriate for high dimensional registration.

We propose a different solution, a new set of multi-modality similarity measures, which we call point similarity measures and can be applied locally [69]. Consequently, the estimation of similarity gradient can easily be numerically estimated and becomes trivial. These measures enable us to perform high-dimensional multi-modality non-rigid registration without any limitation in intensity dependence between the images, while preserving the quality of mono-modality solutions.

1.3 Continuum mechanic non-rigid registration

Continuum mechanics non-rigid registration imitates the real world transformation of deformable materials. The approach was first proposed by Broit [7] in 1981. The transformation is caused by external forces, which act on the deformable media. The deformable media reacts with internal forces, which depend on the material properties and constrain the deformation. The result is an equilibrium state, in which external forces $\mathbf{F}_{ext}$ and internal forces $\mathbf{F}_{int}$ are equal in size and opposite in orientation:

$$\mathbf{F}_{int} = -\mathbf{F}_{ext}. \tag{1.1}$$

In the case of image registration the external forces are obtained such that they tend to improve the matching of the images. The quality of image match is assessed by measuring image similarity, such that the external forces $\mathbf{F} = \mathbf{F}_{ext}$ are normally computed as a gradient of image similarity. Specifically, the external force $\mathbf{F}(\mathbf{x})$, where $\mathbf{x}$ is a world coordinate at which the force acts, is obtained as a gradient of image similarity S with respect to transformation of the point $\mathbf{T}(\mathbf{x})$:

$$\mathbf{F}(\mathbf{x}) = \frac{\partial S}{\partial \mathbf{T}(\mathbf{x})}. \tag{1.2}$$

The relation between the external forces $\mathbf{F}$ and consequent transformation $\mathbf{T}$ is modeled by a spatial deformation model. The spatial deformation model can follow physical properties of deformable media, e.g. elasticity or viscosity, or any other properties that constrain the transformation in a way that suits to the registered anatomy. In general the model must prevent tearing and folding of the images and assure certain smoothness of the transformation.

Continuum mechanic registration can also be performed by minimizing a total system energy E_{system},

$$E_{system} = E_{internal} + E_{external}, \tag{1.3}$$

which also leads to the equilibrium state and transformation that corresponds to the registered images. Here, the external energy $E_{external}$ is defined by the image differences and the internal energy $E_{internal}$ is defined by the spatial deformation model.

A block scheme of continuum mechanic non-rigid registration is shown in Figure 1.3. The registration consists of two basic stages: the estimation of external forces and the spatial deformation model. External forces are computed from a target image A and image $B' = \mathbf{T}B$, which is the source image B transformed with the current transformation $\mathbf{T}$ (initially $\mathbf{T}(\mathbf{x}) = 0$). The obtained forces additionally transform the target image, such that the transformation iteratively converges towards the equilibrium state, which is searched by the registration.

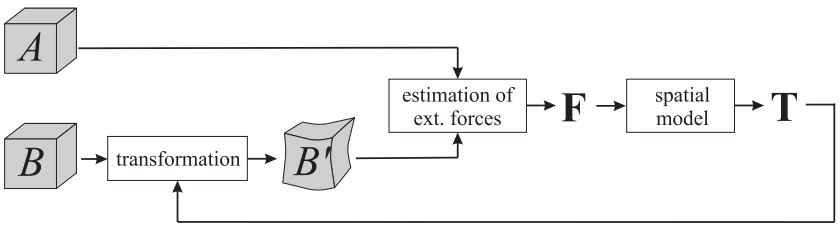


Figure 1.3: *A block scheme of non-rigid registration.*

1.4 Contributions of the thesis

In this thesis we deal with problems related to multi-modality non-rigid registration. Contributions of the thesis are the following ones:

- Point similarity measures, which can measure similarity of individual image points, even in the case of multi-modality images.
- Symmetric image registration approach, which improves registration consistency and registration correctness in comparison to commonly used asymmetric approaches.
- Analysis of convolution based spatial deformation models and development of combined elastic-incremental model.
- Analysis of approaches for evaluation of non-rigid registration and development of a three step evaluation strategy.
- System for multi-modality rigid and non-rigid image registration.

The contributions are described in detail in Chapters 2–5:

Chapter 2: **Point similarity measures**

Point similarity measures are a new set of multi-modality similarity measures, suitable for the multi-modality implementation of non-rigid registration based on continuum mechanics. They are based on the information obtained from the whole images, but can be applied locally. Specifically, they can measure similarity of individual image points. Obviously, they can also be used to measure similarity of image regions of arbitrary size. These measures have several advantages when used for non-rigid registration. In addition to their locality, which enables direct detection of local image discrepancies and improve modeling of spatial deformation properties, they also solve the problem of interpolation artifacts.

Chapter 3: **Symmetric image registration**

Symmetric image registration is a novel approach for improving correctness and consistency of non-rigid registration. It treats both images involved into the registration process in the same manner. Both images may be modeled by spatial deformation models, such that both of them may move and/or deform. Finally, the most distinctive feature of the symmetric registration approach is interaction between the images. Images interact through forces, in accordance with the Newton's third law of motion. Forces on one image reflect in opposing forces on the other image, which forms the basis for symmetry. The result of the interaction is transformation of the images that puts the system into the equilibrium state of minimal energy.

Chapter 4: **Spatial deformation models**

We analyze convolution based spatial deformation models and propose a new combined elastic-incremental model. When elastic and incremental models are used individually, they are faced with certain difficulties. The difficulty of the elastic model is a systematic error, which is caused by restrictions on the overall voxel displacements and makes this model appropriate only for relatively small deformations. On the other hand the incremental model allows large image deformations, but lacks of assuring the anatomical correctness. The proposed combined model reduces the systematic error of the elastic model and is anatomically more correct than the incremental model, and thus improves the registration.

Chapter 5: **Evaluation of multi-modality non-rigid registration**

We analyze techniques for evaluation of non-rigid registration and propose a new evaluation strategy, suitable for registration based on point similarity measures. None of the popular evaluation approaches can detect all kinds of possible misregistrations. The additional problem appears when evaluating multi-modality registration, where only a limited number of features, needed for assessing the actual transformation, can be identified in both of the images. The proposed strategy improves the evaluation of the whole system by separately evaluating each of the two registration stages. This is possible when the registration stages are functionally independent, such as in the case of registration based on point similarity measures.

Chapter 6: **Image registration system**

We present an image registration system, which employs the methods proposed and described in Chapters 2, 3 and 4. The system can be used for rigid as well as for non-rigid registration and is appropriate for various multi-modality and mono-modality tasks. As such it can be used for further research in the field of medical image processing, for medical research and for demonstration of registration capabilities in clinical medicine.

2. Point Similarity Measures

In this chapter we introduce point similarity measures, which we have designed to facilitate performance of multi-modality non-rigid registration. The distinctive feature of point similarity measures is their ability to measure similarity between individual image points, and consequently between appropriately large image regions. Point similarity measures separate the process of similarity computation into two steps. In the first step information of the whole images is used to derive a point similarity function, which is an estimate of intensity dependence between the images. In the second step similarity of individual points is measured using the previously computed point similarity function. The fundamental question concerning point similarity measurement is how to derive a suitable point similarity function, which defines a point similarity measure. We present several point similarity measures, where some of them are derived from other (global) similarity measures, and some of them are designed according to the requirements of multi-modality non-rigid registration. The later ones use knowledge of image intensity distributions and link registration and segmentation. Finally, we analyze the benefits of the proposed measures, compare their point similarity functions and test them as a part of high-dimensional registration system.

2.1 Introduction

High-dimensional non-rigid multi-modality registration detects localized image discrepancies using one of two possible approaches. The first approach measures the similarity of the whole images using well-known multi-modality similarity measures, e.g. mutual information, at applied local image deformations [54, 74]. A weakness of this approach is its high computational cost, which in practice limits the dimensionality and locality of non-rigid registration. The other approach is to measure local image similarity, i.e. the similarity of small image region that gets transformed by a single transformation component. In our high-dimensional case the transformation components equal the displacements of individual image voxels from the initial configuration. This local approach is computationally effective, but requires similarity measures that can operate on such small image regions. Conventional multi-modality similarity measures cannot be used directly

for this purpose, as their statistical nature limits their use to relatively large image regions only. To solve this problem we introduce similarity measures, which can measure similarity of arbitrarily small image regions, including individual image points. Due to their ability to measure similarity of individual image points we call them *point similarity measures*.

2.1.1 Locality of similarity measures

A variety of similarity measures were proposed in order to deal with different requirements of registration, for review see [39, 40]. In general, they can be divided into mono-modality measures, which are appropriate for comparing and registering images of similar intensity characteristics, and multi-modality measures, which can deal with complex image intensity dependencies that are present in multi-modality registration. Similarity measures can also be divided into local and global measures, but this classification is not clear unless the term locality is strictly defined. Local measures can be applied to small image regions, where what is small depends on the application. On the other hand, global techniques require large amount of data and are usually assumed to operate on the whole images. Note that global similarity measures can still be locally sensitive, which means that they can still detect local image differences. Given a local similarity measure, its global equivalent can be derived by summing up over the whole image. On the other hand, global similarity measures may not be able to measure similarity locally, because the performance of global measures decreases by decreasing the region size. The minimal size of the regions that can be used for measuring similarity depends on the nature of the similarity measure in a way that more general measures, with more free parameters, require larger regions.

The simplest similarity measures are based on difference of intensity, e.g. mean absolute difference $MAD(A, B)$ or mean squared difference $MSD(A, B)$,

$$MAD(A, B) = \frac{1}{N} \sum_{v} |i_A(v) - i_B(v)|, \tag{2.1}$$

$$MSD(A, B) = \frac{1}{N} \sum_{v} (i_A(v) - i_B(v))^2, \tag{2.2}$$

where $i_A(v)$ and $i_B(v)$ are image intensities in images A and B at position of voxel v, and N is the number of overlapping voxels. These measures do not tolerate any difference in brightness or contrast, but they do allow the use of arbitrarily small image regions, including individual voxels ($N = 1$).

Measures that can deal with certain intensity differences require larger regions. Correlation based measures, which are the most commonly used mono-modal similarity measures, belong to that group as well. Such measure is the correlation

coefficient $CC(A, B)$ [8],

$$CC(A,B) = \frac{1}{N} \frac{\sum_v (i_A(v) - \overline{i_A})(i_B(v) - \overline{i_B})}{(\sum_v (i_A(v) - \overline{i_A})^2 \sum_v (i_B(v) - \overline{i_B})^2)^{\frac{1}{2}}}. \tag{2.3}$$

Here, $\overline{i_A}$ and $\overline{i_B}$ denote mean image intensity values. Correlation coefficient assumes linear intensity relationship and therefore it can deal with differences in image contrast and brightness. The two unknown parameters of linear intensity relationship (brightness and contrast) require at least three samples (voxels). In practice $3 \times 3 \times 3$ voxel regions or larger regions are used.

The minimum region size requirement becomes much more problematic in the case of cross-modality registration, when the images are acquired using different imaging procedures. In such multi-modality cases the relation between image intensities can be quite complex and is generally not known. Multi-modality similarity measures must be able to deal with such complex relationships. The most widely used multi-modality similarity measures are mutual information $MI(A, B)$ proposed in 1995 by Viola and Wells [83] and Collignon [18], and normalized mutual information $NMI(A, B)$ proposed in 1999 by Studholme et.al. [78],

$$MI(A,B) = H(A) + H(B) - H(A,B), \tag{2.4}$$

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)}. \tag{2.5}$$

Here, $H(A)$, $H(B)$ and $H(A, B)$ are marginal and joint entropies,

$$H(A) = -\sum_{i_A} p(i_A) \log p(i_A), \tag{2.6}$$

$$H(B) = -\sum_{i_B} p(i_B) \log p(i_B), \tag{2.7}$$

$$H(A,B) = -\sum_{\mathbf{i}} p(\mathbf{i}) \log p(\mathbf{i}), \tag{2.8}$$

where $\mathbf{i}$ denotes an intensity pair $[i_A, i_B]$, $p(i_A)$ and $p(i_B)$ are marginal intensity probabilities and $p(\mathbf{i}) = p(i_A, i_B)$ is the joint intensity probability, estimated from the images. Mutual information based measures actually measure statistical dependence between the image intensities, to estimate how much one image tells about the other one. But when the image regions used to estimate the joint distribution are small, its statistical significance is low and the similarity measures poorly represent actual image correspondence.

Various solutions have been proposed in order to improve locality of multi-modality measures. The most obvious solution is intensity binning, i.e. decreasing the number of intensity bins [47]. This directly improves the estimation of joint

distribution, but on the other hand it lowers the intensity sensitivity, e.g. different features that are represented by similar intensity values can be grouped into the same bin. The number of bins can also be reduced by resorting to one dimensional statistics, e.g. $H(A-B)$ or $H(A+B)$ [10], which also improves the locality. Similar to the binning is Parzen window estimation of intensity distributions [58]. In the case of discrete intensity space it can be efficiently implemented as post-filtering of normalized joint histogram. This method lowers intensity sensitivity as well.

Another approach for improving the locality encompasses methods that use prior knowledge. In the case of multi-modality measures such knowledge is usually given in a form of a joint distribution p_{prior} of correctly registered images. Such a solution was used in the log likelihood similarity measure proposed by Leventon and Grimson [44]. Of course, it is also possible to use a combination of both, a prior probability p_{prior} that is given in advance, and probability p_{image}, which is estimated from the images:

$$p = \lambda p_{image} + (1-\lambda) p_{prior}, \tag{2.9}$$

where λ is a weighting parameter. This enables additional reduction of region size [47]. As correct prior probability is rarely known, Maintz et al. proposed to use a global distribution instead of prior one [52]. Their local measure is based on global conditional probability. Joint probability estimated from smaller image regions instead of the whole images can also be used in place of prior information, as demonstrated in [46].

To solve the problem of locality we propose a new group of similarity measures, which we call *point similarity measures*. They are based on the information obtained from the whole images but are applied locally, on arbitrarily small image regions, including individual image voxels/points.

2.2 The basic principle of point similarity measures

Point similarity measures divide the process of similarity measurement into two steps. In the first step a point similarity function $f(\mathbf{i})$ is obtained. Point similarity function is an estimate of intensity dependence between the images A and B. The information used in this step is not limited to local image region. It may include information extracted from the whole images, as well as potential prior knowledge. In the second step actual measurement of similarity $S(\mathbf{x}_1, \mathbf{x}_2)$ between two image points $A(\mathbf{x}_1)$ and $B(\mathbf{x}_2)$ takes place. Thus, a point at world coordinate $\mathbf{x}_1$ in image A is compared with point at world coordinate $\mathbf{x}_2$ in image B. Here, the point similarity function $f(\mathbf{i})$ estimated in the first step is used to provide the necessary information about image intensity dependence, such that the similarity $S(\mathbf{x}_1, \mathbf{x}_2)$ equals the value of point similarity function $f(\mathbf{i}) = f(i_A, i_B)$ at given

intensities of the compared image points, $i_A(\mathbf{x}_1)$ and $i_B(\mathbf{x}_2)$.

$$S(\mathbf{x}_1, \mathbf{x}_2) = f(i_A(\mathbf{x}_1), i_B(\mathbf{x}_2)), \tag{2.10}$$

Let us illustrate the principle of point similarity measures on a mono-modality similarity measure MAD, see Eq. 2.1. In the first step the point similarity function is defined, in this case by the prior knowledge of mono-modality intensity distributions:

$$f_{MAD}(\mathbf{i}) = -|i_A - i_B|. \tag{2.11}$$

The negative sign is used to obtain higher similarity for more similar images. In the second step, when the point similarity is defined, measuring of point similarities takes place. This step is simple and does not depend on the type of the point similarity measure. For the illustration see Fig. 2.1. Any two points $A(\mathbf{x}_1)$ and $B(\mathbf{x}_2)$ form a point pair with its characteristic intensity pair $\mathbf{i} = [i_A(\mathbf{x}_1), i_B(\mathbf{x}_2)]$. The value of point similarity function at this intensity pair $f(\mathbf{i})$ represents the similarity of these two points.

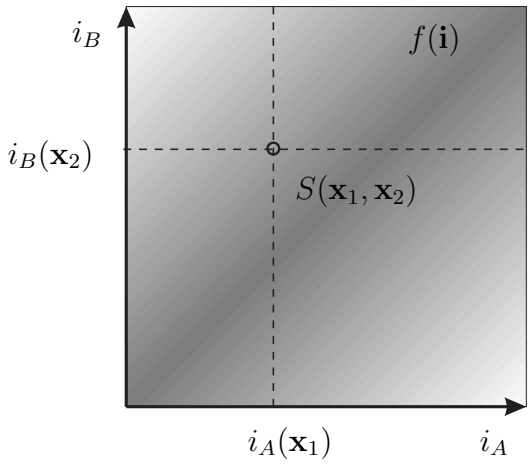


Figure 2.1: *An illustration of point similarity measurement. The similarity* $S(\mathbf{x}1, \mathbf{x}2)$ *of two points* $A(\mathbf{x}_1)$ *and* $B(\mathbf{x}_2)$ *is obtained from a similarity function* $f(\mathbf{i})$*, by pointing to a certain value by a corresponding intensity pair* $\mathbf{i} = [i_A(\mathbf{x}_1), i_B(\mathbf{x}_2)]$*. In this example the similarity function corresponds to MAD similarity measure, such that* $f(\mathbf{i}) = -|i_A - i_B|$ *(darker color represents higher similarity).*

The only difference between different point similarity measures is in the first step, i.e., in the derivation of point similarity function $f(\mathbf{i})$. Instead of using prior knowledge, as it is in the case of MAD measure, it can also be derived from the images that are being registered. Numerous approaches can be used to derive the point similarity function from the images. Because every improvement of some point similarity $S(\mathbf{x}_1, \mathbf{x}_2)$ also reflects in the improvement of some

global similarity $S_G(A, B)$ between the whole images, it is possible to derive point similarity functions from global similarity measures. The derivation is the most straightforward when global similarity is computed as an average (or sum) of local contributions $S(v)$, defined for each individual voxel v:

$$S_G(A, B) = \frac{1}{N} \sum_v S(v) = \overline{S(v)}, \tag{2.12}$$

where N is number of voxels in the image. In this case $S(v)$ is point similarity for voxel v, i.e. a similarity between a point $A(\mathbf{x}(v))$ and an overlapping point $B(\mathbf{x}(v))$:

$$S(v) = S(\mathbf{x}(v), \mathbf{x}(v)). \tag{2.13}$$

For example, a mono-modality measure $S_G(A, B) = -MAD(A, B)$ [1] is defined as a spatial average of voxel contributions (2.1). Consequently, the corresponding point similarity is $S(\mathbf{x}_1, \mathbf{x}_2) = -|i_A(\mathbf{x}_1) - i_B(\mathbf{x}_2)|$, and the point similarity function is therefore $f(\mathbf{i}) = -|i_A - i_B|$, see Fig. 2.1.

Whenever a global similarity S_G can be computed by averaging point similarities S as noted in 2.12, then it can also be computed by weighted averaging of values of the point similarity function over all intensity pairs, where weights are the joint intensity probabilities:

$$S_G(A, B) = \sum_{\mathbf{i}} p(\mathbf{i}) f(\mathbf{i}). \tag{2.14}$$

To prove this relation let us rewrite Eq.2.12 using the Eq.2.12:

$$\begin{aligned} S_G(A, B) &= \frac{1}{N} \sum_v f(\mathbf{i}(v)) \\ &= \frac{1}{N} \sum_v \sum_{\mathbf{j}} f(\mathbf{j}) \delta(\mathbf{i}(v) - \mathbf{j}) \\ &= \sum_{\mathbf{j}} f(\mathbf{j}) \frac{1}{N} \sum_v \delta(\mathbf{i}(v) - \mathbf{j}) \\ &= \sum_{\mathbf{j}} f(\mathbf{j}) \frac{N_{\mathbf{j}}}{N} \\ &= \sum_{\mathbf{j}} f(\mathbf{j}) p(j). \qquad \square \end{aligned} \tag{2.15}$$

Here, δ is a Dirac's delta function and $N_{\mathbf{j}}$ stands for the number of voxels with corresponding joint intensity pair $\mathbf{j}$. Furthermore, it is considered that joint intensity distribution p can be computed by normalizing joint histogram, such that $p(\mathbf{i}) = \frac{N_{\mathbf{i}}}{N}$.

[1] a negative sign is used to get higher similarity for better image match

Point similarity measures can be derived from any intensity based similarity measure, if not analytically they can still be obtained numerically. Let $\mathbf{T}$ denote a local transformation that transforms image B to image B' and therefore also changes the image point pairs, such that intensity pair $\mathbf{i}(v)$ at point v changes to $\mathbf{i}'(v)$:

$$\mathbf{T}: \ \mathbf{i}(v) \longrightarrow \mathbf{i}'(v), \ B \longrightarrow B'. \tag{2.16}$$

These local changes must be detected by point similarity measure S, such that the new point similarity of voxel v is

$$S'(v) = S(\mathbf{i}'(v)). \tag{2.17}$$

Assuming that larger similarity means better match, the increase of point similarities $S(v)$ must increase some global similarity $S_G(A, B)$:

$$\sum_n S'(v_n) > \sum_n S(v_n) \Longrightarrow S_G(A, B') > S_G(A, B), \tag{2.18}$$

where the summations run over all image voxels. If the transformation $\mathbf{T}$ changes the image match at only one image voxel v_m and point similarities of all other image voxels remain unchanged, $S(v_n) = S'(v_n)$, $n \neq m$, then eq.(2.18) can be rewritten:

$$\sum_{n;n\neq m} S(v_n) + S'(v_m) > \sum_{n;n\neq m} S(v_n) + S(v_m) \Longrightarrow S_G(A, B') > S_G(A, B), \tag{2.19}$$

and therefore

$$S'(v_m) > S(v_m) \Longrightarrow S_G(A, B') > S_G(A, B). \tag{2.20}$$

Hence, the improvement of image correspondence at a single image point v_m improves the global image correspondence as well. This relation can be used for deriving point similarity measures from global intensity based similarity measures. Let us imagine we can find a local transformation $\mathbf{T}$, which changes image match in a single point v_m, e.g change the intensity pair $\mathbf{i}$ at that point. Then the change of point similarity $S(v_m)$ must correspond to the change of global similarity $S_G(A, B)$. One way to define point similarity measure is therefore the following:

$$S'(v_m) = f(\mathbf{i}'(v_m)) = C_1 \cdot S_G(A, B') + C_0; \tag{2.21}$$

where constants C_1 and C_0 can be chosen arbitrarily, whereas the sign of C_1 should be such that larger point similarity means better image match[2]. Two point measures that differ only in constants C_1 and C_0, e.g. $S_1(\mathbf{i}'(v_m)) = 2 \cdot S_G(A, B')$ and $S_2(\mathbf{i}'(v_m)) = S_G(A, B') - S_G(A, B)$, are equivalent and they both reach the optimum at the same local image transformation.

[2] C_1 is negative if lower value of S_G means better correspondence.

Thus, point similarity function $f(\mathbf{i})$ can always be estimated from global similarity measure S_G, either analytically as shown in Eq.(2.12), or numerically by changing image correspondence at single image point. This also holds for multi-modality similarity measures. In such cases point similarity function has to adopt to actual image properties, e.g. they may depend on actual image intensity distributions. In the following section we analytically derive point similarity measures from global multi-modality measures.

2.3 Multi-modality point similarity measures

The difficulty in measuring multi-modality similarity as compared to mono-modality similarity measurement is that the intensity dependence between the images is complex and in general unknown. The estimation of such complex intensity dependence, which is a part of similarity measurement, requires a sufficiently large amount of data and thus cannot be performed on the basis of small image regions. Point similarity measures solve this problem by separating the process of similarity measurement into two steps. In the first step the similarity function $f(\mathbf{i})$ is obtained, using the information of the whole images. In the second step the similarity of two points, i.e. point similarity, is measured using the similarity function obtained in the first step, as described in previous section and illustrated in Fig. 2.1. The fundamental step in measuring multi-modality point similarities is therefore derivation of point similarity function. In this section we derive point similarity functions from global multi-modality similarity measures.

2.3.1 Entropy based point similarity measure

Joint entropy $H(A, B)$ forms the basis of many multi-modality similarity measures. It is the most informative part of mutual information, and can be independently used as a global multi-modality similarity measure [19]. In this section we derive a point similarity measure from a negative joint entropy $-H(A, B)$, such that higher similarity corresponds to better image match.

Let us rewrite (2.8) in the following form,

$$-H(A,B) = \sum_{\mathbf{j}} p(\mathbf{j}) \log p(\mathbf{j}) = \sum_{\mathbf{j}} \frac{N_{\mathbf{j}}}{N} \log p(\mathbf{j}) = \frac{1}{N} \sum_{\mathbf{j}} N_{\mathbf{j}} \log p(\mathbf{j}) \qquad (2.22)$$

where $\mathbf{j} \in \{\mathbf{i}\}$ runs over all intensity pairs, $N_{\mathbf{j}}$ is the number of occurrences of intensity pair $\mathbf{j}$ and N is the total number of intensity pairs in the image, which equals the number of overlapping image voxels. $N_{\mathbf{j}}$ can be computed by counting image voxels v with intensity pair $\mathbf{i}(v) = \mathbf{j}$ over the whole images:

$$N_{\mathbf{j}} = \sum_{v} \delta(\mathbf{i}(v) - \mathbf{j}). \qquad (2.23)$$

Here the summation runs over all image voxels v and δ is a Dirac's delta function. The Eq. 2.22 can thus be rewritten as follows:

$$-H(A,B) = \frac{1}{N}\sum_{\mathbf{j}}\sum_{v}\delta(\mathbf{i}(v)-\mathbf{j})\log p(\mathbf{j}). \tag{2.24}$$

As $\delta(\mathbf{i}(v)-\mathbf{j})$ differs from zero only when $\mathbf{i}(v)=\mathbf{j}$, it is possible to omit the summation over the intensities:

$$-H(A,B) = \frac{1}{N}\sum_{v}\log p(\mathbf{i}(v)). \tag{2.25}$$

Note that the final summation is taken over the spatial image coordinates instead of intensities. Thus, global similarity $-H(A,B)$ can be treated as an average of local contributions, defined for each voxel v.

$$-H(A,B) = \frac{1}{N}\sum_{v} S_H(v), \tag{2.26}$$

$$S_H(v) = \log p(\mathbf{i}(v)). \tag{2.27}$$

When transformation $\mathbf{T}$ changes image B to B', the contribution of voxel v to global joint entropy changes as well:

$$S'_H(v) = \log p'(\mathbf{i}'(v)). \tag{2.28}$$

There are two sources of the change. First, the transformation $\mathbf{T}$ changes the matching of voxel v, which changes the corresponding intensity pair from $\mathbf{i}(v)$ to $\mathbf{i}'(v)$, and second, the transformation also changes the joint distribution p to p'. But the change of joint distribution does not depend only on the transformation of one voxel, but on the transformation of the whole image, which is not known until all local transformations (voxels displacements) are obtained. So, let us express $p'(\mathbf{i}'(v))$ using Taylor expansion:

$$S'_H(v) = \log\left(p\left(\mathbf{i}'(v)\right) + \frac{\partial p(\mathbf{i}'(v))}{\partial \mathbf{T}}\mathbf{T} + \frac{1}{2}\frac{\partial^2 p(\mathbf{i}'(v))}{\partial \mathbf{T}^2}\mathbf{T}^2 + ...\right) \tag{2.29}$$

As the transformation $\mathbf{T}$ that correspond to one iteration of image registration is supposed to be small, the terms above the zero order are small as well and can be neglected. Only the the zero order term remains such that the similarity can be estimated as:

$$S'_H(v) = \log p\left(\mathbf{i}'(v)\right), \tag{2.30}$$

The obtained estimation of local contribution to the global similarity can be used as a point similarity measure, where the corresponding similarity function is

$$f_H(\mathbf{i}) = \log p(\mathbf{i}). \tag{2.31}$$

Notice that although the similarity is measured for individual image points, it still requires the estimation of joint probabilities $p(\mathbf{i})$ from the whole images. The obtained entropy based point similarities S_H then represent the amount of information provided by the intensity pair $\mathbf{i}$.

Because $p(\mathbf{i})$ is used instead of unknown $p'(\mathbf{i})$, point similarity function corresponds to joint entropy only for small transformations $\mathbf{T}$. However, even when this is not he case, the point similarity measure S_H remains suitable for measuring multi-modality image similarity.

2.3.2 MI based point similarity measure

The same approach as used for derivation of entropy based point similarity measure S_H can also be used for the derivation of mutual information based point similarity measure S_{MI}. Mutual information (2.4) can be computed as follows:

$$MI = \sum_{\mathbf{j}} p(\mathbf{j}) \log \left(\frac{p\,(\mathbf{j})}{p\,(j_A)\,p\,(j_B)} \right). \tag{2.32}$$

where $\mathbf{j} = [j_A, j_B] \in \{\mathbf{i}\}$ runs over all intensity pairs. This equation can be rewritten in the following form:

$$MI = \frac{1}{N} \sum_{\mathbf{j}} N_{\mathbf{j}} \log \left(\frac{p\,(\mathbf{j})}{p\,(j_A)\,p\,(j_B)} \right). \tag{2.33}$$

Here $N_{\mathbf{j}}$ again denotes the number of occurrences of intensity pair $\mathbf{j}$ and N is the total number of intensity pairs in the image, which equals the number of overlapping image voxels. Using Eq. 2.23 the Eq. 2.33 can thus be rewritten as

$$MI = \frac{1}{N} \sum_{\mathbf{j}} \sum_{v} \delta(\mathbf{i}(v) - \mathbf{j}) \log \left(\frac{p\,(\mathbf{j})}{p\,(j_A)\,p\,(j_B)} \right). \tag{2.34}$$

The delta function differs from zero only when $\mathbf{i}(v) = \mathbf{j}$, such that the summation over the intensities can be omitted:

$$MI = \frac{1}{N} \sum_{v} \log \left(\frac{p(\mathbf{i}(v))}{p(i_A(v))p(i_B(v))} \right), \tag{2.35}$$

Note that the final summation is taken over the spatial image coordinates and not over the intensities. Thus, the global similarity MI can be treated as an average of point similarities $S_{MI}(v)$, defined for each voxel v.

$$MI = \frac{1}{N} \sum_{v} S_{MI}(v), \tag{2.36}$$

$$S_{MI}(v) = \log\left(\frac{p(\mathbf{i}(v))}{p(i_A(v))p(i_B(v))}\right). \tag{2.37}$$

The mutual information based point similarity function is therefore the following:

$$f_{MI}(\mathbf{i}) = \log\left(\frac{p(\mathbf{i})}{p(i_A)p(i_B)}\right). \tag{2.38}$$

When image B gets transformed with transformation $\mathbf{T}$ the joint intensity distribution $p(\mathbf{i})$ changes and, in general, this also changes the similarity function $f_{MI}(\mathbf{i})$. Anyway, the point similarity function is always an estimate of the same intensity dependence, i.e. intensity dependence of the registered images. Thus, the similarity can be measured even if the original point similarity function is used instead of the new one. However, note that the global meaning of mutual information is preserved only if transformation $\mathbf{T}$ is small.

2.3.3 Other point similarity measures

Point similarity measures can be directly derived from almost any intensity based similarity measure. For example, a point similarity measure that corresponds to the energy similarity measure [11]:

$$P = \sum_{\mathbf{i}} p^2(\mathbf{i}), \tag{2.39}$$

has a point similarity function f_P that equals the estimated joint intensity distribution:

$$f_P(\mathbf{i}) = p(\mathbf{i}). \tag{2.40}$$

Point similarity measures can also be defined by conditional probabilities [52, 26]:

$$f_{PC}(\mathbf{i}) = p(i_A|i_B). \tag{2.41}$$

By additionally applying a log function, point similarity becomes related to information instead to probability. Thus, the log function applied on the joint probability $p(\mathbf{i}) = f_P(\mathbf{i})$ results in the entropy based point similarity function $f_H = \log p(\mathbf{i})$, see Section 2.3.1. Similarly, the information can also be measured using conditional probability:

$$f_{HC}(\mathbf{i}) = \log p(i_A|i_B). \tag{2.42}$$

All these point similarity measures are based on statistics, specifically on the estimated global joint intensity distribution, which is used for estimation of intensity dependence. The intensity dependence is better estimated when images are better matched and cannot be exactly determined before the images are registered. However, the estimation can be improved by using additional knowledge. One of the possibilities is to use additional knowledge about intensity distributions, which links point similarity measures and therefore registration with segmentation.

2.4 Point similarity measures and segmentation

Some of the intensities that are present in the images are closely related and represent the same tissue type. If their close relation is taken into account, an improvement in the registration can be expected. Finding such relations between the intensities is called image segmentation and allows registration to operate directly with tissue types instead of image intensities. Thus, registration and segmentation are related. Segmentation can be performed by registering image to an anatomical atlas, while registration can gain from better biological image representation, obtained by segmentation. In this section we propose a method that integrates both tasks, such that they can be performed simultaneously. Due to the additional information, which can be used for each of the tasks, it is expected that such simultaneous operation should improve registration as well as segmentation.

Both, segmentation and registration, operate on the intensity distributions. The first one groups image intensities into classes, which represent different tissue types, while the second one uses joint distributions for computing multi-modality point similarities. Thus, to understand point similarity measures and relate them to segmentation, it is important to understand joint intensity distributions and their relation to the quality of image match.

2.4.1 Joint intensity distributions

Joint intensity distribution is a quantitative representation of the intensity relationship at certain image alignment, and changes during the registration [76]. Only the joint distribution of correctly matched images corresponds to the actual intensity relationship. However, the actual intensity relationship is not known until the images are correctly registered. In this section we analyze how joint intensity distribution changes with respect to the degree of image misalignment, and how it is related to image segmentation.

Imagine we have two simple images A and B representing the same object, see Fig. 2.2. Let each image consist of only two intensity values, i_{1A} and i_{2A} for image A, and i_{1B} and i_{2B} for image B, where i_{1A} corresponds to i_{1B} and i_{2A} corresponds to i_{2B}. When images are correctly registered, the joint distribution consists of only two extrema, at intensity pairs $[i_{1A}, i_{1B}]$ and $[i_{2A}, i_{2B}]$, because the intensity regions perfectly overlap. Let us call these intensity pairs true intensity pairs, as they correspond to the correctly matched image regions. When images do not overlap exactly, additional 'false' intensity pairs appear, in our case $[i_{1A}, i_{2B}]$ and/or $[i_{2A}, i_{1B}]$. Probabilities of these true and false intensity pairs depend on the size of overlapping regions.

Real images are corrupted by noise and intensity distortion, and so each region

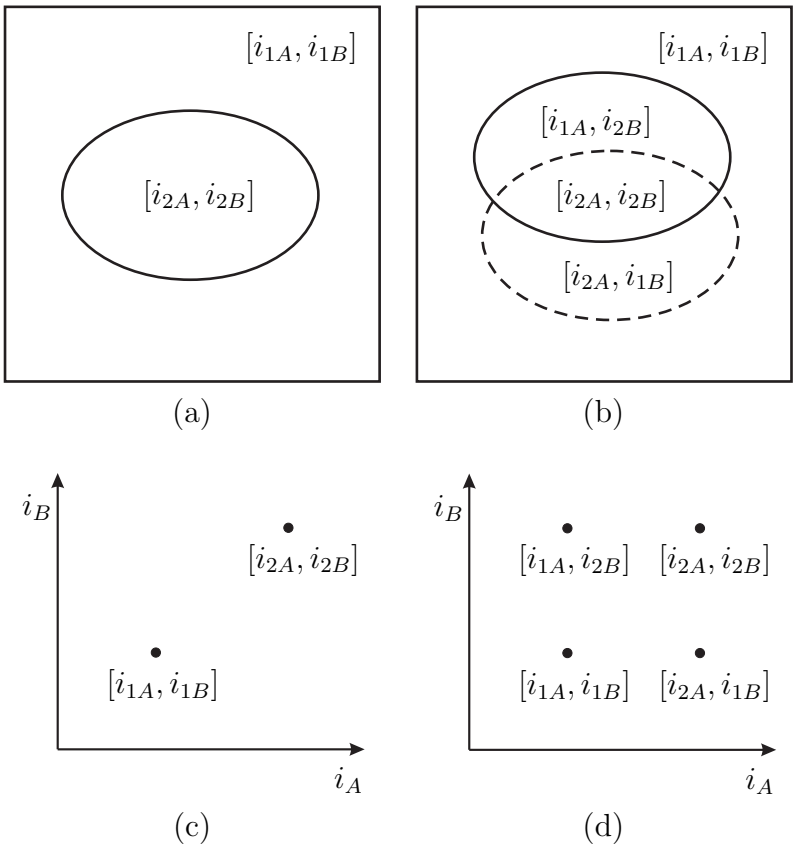


Figure 2.2: *Joint distribution (c) of registered images (a) and joint distribution (d) of misregistered images (b).*

overlap is represented by numerous intensity pairs that form an intensity class. Therefore, true and false intensity pairs become true and false intensity classes. Each class represents a tissue type pair. Furthermore, real images consist also of some intensities that do not belong to any of the mentioned intensity classes. In case of 3D anatomical medical images most of them appear on the transitions between different tissues and belong to partial volume voxels. In joint distribution they are positioned within a rectangle defined by corresponding intensity classes, see examples in Fig. 2.3.

Point similarity functions presented in Section 2.3 are closely related to the joint intensity distributions. However, the similarity function should model intensity dependence of correctly registered images and should have large values only for intensity pairs that belong to true classes, while similarity values for false classes should be minimal. Furthermore, intensity pairs from true classes should all produce comparable similarity values. Unfortunately similarity functions differ from such ideal form. A higher level of image misalignment reflects in

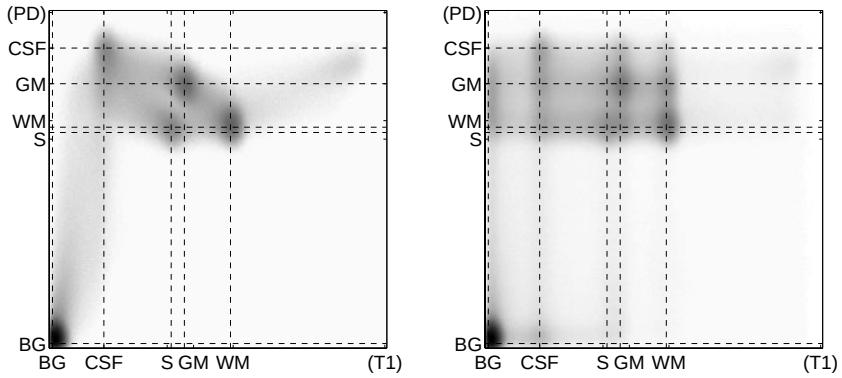


Figure 2.3: *Examples of joint distributions for MRI T1 and PD medical images of head: registered images (a) and not registered images (b). Dashed lines correspond to mean intensity values of tissue types: cerebrospinal fluid (CSF), gray matter (GM), white matter (WM), surrounding tissues (S), and background (BG). Partial volume voxels are positioned inside rectangles defined by these lines.*

higher estimated probabilities for false classes, while probabilities for true classes decrease. The described similarity functions do not discriminate between true and false classes and are directly related to the estimated probabilities. Thus, if the level of misalignment is too high, similarities for false classes become larger than similarities for true classes, and non-rigid registration in such case does not converge to the correct image match. Moreover, similarities of intensity pairs also depend on the amount of tissues that form these intensities. This may result in differences in their registration speed and registration correctness. To solve these problems we develop a segmentation based point similarity measure, which combines registration with segmentation.

2.4.2 Segmentation based point similarity measure

By modeling the intensity classes, the point similarity function can be estimated as a probability of an intensity pair belonging to one of the true classes. Such an estimatie improves two similarity function properties. Firstly, it equalizes similarities among correctly matched tissues in the sense that probability that an intensity pair belongs to a certain true class is not related to the amount of corresponding tissue. Secondly, such a similarity measure lowers similarities for incorrectly matched regions, as knowledge of intensity distribution can be used to distinguish between true and false classes, see Subsection 2.4.1 and Figs. 2.2 and 2.3.

The basic idea of the segmentation based point similarity measure is to define

its point similarity function as a probability of certain intensity pair to belong to one of the true classes:

$$f_S(\mathbf{i}) = p(\mathcal{C}_T|\mathbf{i}) = \sum_m p(C_m|\mathbf{i})p(\mathcal{C}_T|C_m). \tag{2.43}$$

Here, C_m denotes intensity class, which represents certain tissue type pair, and $\mathcal{C}_T$ is a set of all true intensity classes. Remember that true classes are classes which correspond to the same tissue in both images. The first factor in (2.43) therefore represents probability of m-th class at a given intensity pair $\mathbf{i}$. The second factor is a probability that certain class C_m is a true class and therefore belongs to $\mathcal{C}_T$.

The first step towards segmentation based similarity measure is modeling of joint distribution, which consists of several classes, each of them representing a different tissue type pair. Joint distribution $p(\mathbf{i})$ is thus a weighted sum of individual class distributions $p(\mathbf{i}|C_m)$,

$$p(\mathbf{i}) = \sum_m p(\mathbf{i}, C_m) = \sum_m p(\mathbf{i}|C_m)p(C_m). \tag{2.44}$$

The class distributions can be obtained by segmenting the images into the tissue types and then estimating the probability distribution of each tissue type pair. This approach requires pre-segmentation of images, which is not an easy task. Various image segmentation methods can be used, for review see [16, 61]. Note that image segmentation is in close relation with modeling intensity distribution. Not only that joint distribution can be modeled by using segmentation, the segmentation is also often based on modeling intensity distribution.

We propose another approach, which models joint distribution directly and is equivalent to simultaneous segmentation of both images. Because images are not yet registered the number of intensity classes is initially not known and precise modeling is difficult, as most standard approaches cannot be directly used. Our approach automatically detects how many intensity classes exist and is computationally efficient, as it models only the required single-tissue intensities, excluding mixtures that appear in partial volume voxels. Assuming Gaussian intensity distribution within each tissue type, each class C_m can be modeled by a 2D Gaussian function with mean value μ_m and covariance matrix $\mathbf{\Sigma}_m$. Low intensities, e.g. background, which are approximately Rayleigh distributed [37] can be modeled by a Gaussian model sufficiently well, because the significant difference appears only at lowest intensities, where probability of other tissues is low. Let us assume that classes are far enough from each other to achieve dominance of class C_m in its neighborhood $\mathcal{O}_m$, such that contributions of all other classes can be neglected. Classes that are merging can be modeled together as a single class and separated in later registration steps, when their overlap decreases. Number of classes M, their mean values μ_m and joint intensity distribution maxima a_m

can be estimated by an exhaustive search for maxima in joint intensity distribution. When maximum is found, its position is used as a class mean value μ, while the value itself is used as an amplitude of probability $p(\mathbf{i}, C_m)$. Probabilities of intensity pairs in $\mathcal{O}_m$ can be approximated by

$$p(\mathbf{i})|_{\mathbf{i}\in\mathcal{O}_m} \approx p(\mathbf{i}, C_m) = a_m \exp\left(-\frac{1}{2}(\mathbf{i}-\mu_m)^T\mathbf{\Sigma}_m^{-1}(\mathbf{i}-\mu_m)\right) \;\; ; \quad m = 1\ldots M. \tag{2.45}$$

By taking a logarithm of (2.45) we get

$$2\ln\left(\frac{a_m}{p(\mathbf{i})}\right) = (\mathbf{i}-\mu_m)^T\mathbf{\Sigma}_m^{-1}(\mathbf{i}-\mu_m) \quad ; \quad \mathbf{i}\in\mathcal{O}_m \tag{2.46}$$

$$\mathbf{\Sigma}_m^{-1} = \begin{bmatrix} u_{11} & u_{12} \\ u_{12} & u_{22} \end{bmatrix} \tag{2.47}$$

$$2\ln\left(\frac{a_m}{p(\mathbf{i})}\right) = u_{11}(i_A-\mu_{mA})^2 + 2u_{12}(i_A-\mu_{mA})(i_B-\mu_{mB}) + u_{22}(i_B-\mu_{mB})^2 \tag{2.48}$$

which can be solved for $\mathbf{\Sigma}_m^{-1}$ using least squares method for all intensity pairs $\mathbf{i}$ in the neighborhood $\mathcal{O}_m$. The covariance matrices $\mathbf{\Sigma}_m$ can then be used to estimate the class a priori probabilities $p(C_m)$.

$$\begin{aligned} p(C_m) &= \int p(\mathbf{i}, C_m)d\mathbf{i} = \\ &= \int a_m \exp\left(-\frac{1}{2}(\mathbf{i}-\mu_m)^T\mathbf{\Sigma}_m^{-1}(\mathbf{i}-\mu_m)\right) d\mathbf{i} \\ &= a_m 2\pi|\mathbf{\Sigma}_m| \quad ; \quad m = 1..M \end{aligned} \tag{2.49}$$

Theoretically the sum of all a priori probabilities $p(C_m)$, $m = 1...M$, should be 1. In reality this is seldom the case even if all class parameters are estimated absolutely correct, due to the fact that some intensity pairs with low joint probabilities do not belong to any of the estimated classes. The majority of such intensity pairs represent partial volume (PV) voxels. The correct position of PV voxels in joint distribution is hard to predict, as they can be positioned anywhere in the rectangle defined by the true classes, see Fig. 2.3. Currently we are not dealing with partial volume voxels explicitly. We simply model them with additional class C_0 with uniform distribution:

$$p(\mathbf{i}, C_0) = \varepsilon. \tag{2.50}$$

The selection of ε is not critical and we set it to $1/N$ where N is a total number of overlapping voxels.

A posterior probability $p(C_m|\mathbf{i})$ of class C_m, which shows the chance that certain intensity pair $\mathbf{i}$ belongs to a particular class C_m, is according to Bayes

rule

$$p(C_m|\mathbf{i}) = \frac{p(\mathbf{i}, C_m)}{\sum_{l=0}^{M} p(\mathbf{i}, C_l)}. \tag{2.51}$$

This concludes the modeling of joint intensity distribution. Other modeling or segmentation approaches can be used as well. The described method is used because of its low computational cost at sufficient accuracy. The obtained model depends on correctness of match and improves during the registration. Thus, image registration and image segmentation/modeling are performed simultaneously.

After joint distribution model is obtained, probabilities $p(\mathcal{C}_T|C_m)$ that certain class C_m is a true class must be estimated in order to determine point similarity function (2.43). The set $\mathcal{C}_T$ includes all the true classes. We cannot know exactly which classes are true classes, but using knowledge of joint distributions, it is possible to estimate this probability for each of the classes. Let us assume that each tissue type has a unique intensity representation with mean value μ. Therefore, among all maxima positioned at the same intensity of image A (or image B) only one can belong to the set of true classes $\mathcal{C}_T$. Let a set of classes $\mathcal{C}_{\mu A}$ consist of all classes with the same mean value μ_A (we allow a difference of one standard deviation), then it is expected that classes $C_m \in \mathcal{C}_{\mu A}$ with higher probabilities $p(C_m)$ are more likely to be true classes. Therefore, we can estimate probability $p_A(\mathcal{C}_T|C_m)$ of class C_m being a true class according to image A as follows:

$$p_A(\mathcal{C}_T|C_m) = \frac{p(C_m)}{\sum_{C_l \in \mathcal{C}_{\mu A}} p(C_l)}; \quad C_m \in \mathcal{C}_{\mu A}. \tag{2.52}$$

Such a probability estimation of a certain class being a true class is not sufficient. For example, when one of the tissues dominates, all classes with μ_B that correspond to that tissue could have the highest probability $p_A(\mathcal{C}_T|C_m)$, although it is expected that only one of them is a true class. This problem can be resolved by using probability $p_B(\mathcal{C}_T|C_m)$ of class C_m being a true class according to image B. Estimation of this probability is equivalent to the estimation of $p_A(\mathcal{C}_T|C_m)$:

$$p_B(\mathcal{C}_T|C_m) = \frac{p(C_m)}{\sum_{C_l \in \mathcal{C}_{\mu B}} p(C_l)}; \quad C_m \in \mathcal{C}_{\mu B}, \tag{2.53}$$

where a set $\mathcal{C}_{\mu B}$ comprises all the classes with the same μ_B. Class C_m can be assumed to be a true class only if it is a true class according to image A as well as to image B. The estimates $p_A(\mathcal{C}_T|C_m)$ and $p_B(\mathcal{C}_T|C_m)$ are based on different observations so they can be considered independent. Therefore, the final estimate of probability that class C_m is a true class can be obtained as a product of both probabilities:

$$p(\mathcal{C}_T|C_m) = p_A(\mathcal{C}_T|C_m) \cdot p_B(\mathcal{C}_T|C_m). \tag{2.54}$$

If two or more tissue types have the same intensity representation μ_A or μ_B our presumption is incorrect. Let us suppose there are ω tissue types with the same μ_A or μ_B and therefore, there should also be ω corresponding true classes. Our estimated probabilities $p_A(\mathcal{C}_T|C_m)$ or $p_B(\mathcal{C}_S|C_m)$ of these classes are reduced, in general ω times. Nevertheless, ratios between probabilities $p(\mathcal{C}_T|C_m)$ of these true classes and their competitive false classes remain in proportion and so the registration should still tend to correctly minimize the false classes, although it may require more registration iterations. Of course, tissues with the same μ_A or μ_B cannot be distinguished.

Once the required probabilities are estimated (i.e. $p(C_m|\mathbf{i})$ and $p(\mathcal{C}_T|C_m)$), segmentation based point similarity function $f_S(\mathbf{i})$ can be computed using (2.43).

Note that segmentation based point similarity measure $S_S(\mathbf{x}_1, \mathbf{x}_2) = f_S(i_A(\mathbf{x}_1), i_B(\mathbf{x}_2))$ requires that the image data form intensity classes, which can be modeled by Gaussian functions. This is a reasonable assumption and does not need to be exactly true in practice. Due to such modeling of joint distribution measure S_S does not model partial volume voxels. Matching of partial volume voxels is difficult as their intensity relations are not known. Linear intensity relation is often suggested, but in reality it may not be valid due to the multimodality. Measure S_S does not assume any relation between partial volume voxel intensities. It only matches pure tissues and when a correct spatial deformation model is used, partial volume voxels are expected to match correctly as well.

2.4.3 Point similarity measures without segmentation

In some cases image intensity distribution does not form intensity classes. Intensities of medical images form classes in case of 3D anatomical imaging techniques, but in other cases classes may not be formed. Furthermore, even if classes exist they may be difficult to model when images include large intensity variations within tissue types, e.g. intensity inhomogeneity. Functional techniques do not form intensity classes at all. Consequently, matching of such images using previously described segmentation based measure S_S is not expected to produce good results. Nevertheless, some intensity pairs always represent correct match and some other intensity pairs represent incorrect match. Every intensity of source image tends to match some intensities of target image, and vice versa. Therefore, each intensity pair can be treated as its own intensity class $C_{\mathbf{i}}$ with probability $p(C_{\mathbf{i}}) = p(\mathbf{i})$, mean value $\mu = \mathbf{i}$ and $p(C_{\mathbf{i}}|\mathbf{i}) = \delta(\mathbf{i})$, where $\delta(\mathbf{i})$ is a Dirac's delta function. If these classes are used the same way as in segmentation based similarity measure, a new point similarity measure S_U is obtained. Its point similarity function f_U can be estimated as follows:

$$p_A(\mathcal{C}_T|\mathbf{i}) = \frac{p(\mathbf{i})}{p(i_A)} = p(i_B|i_A), \quad p_B(\mathcal{C}_T|\mathbf{i}) = \frac{p(\mathbf{i})}{p(i_B)} = p(i_A|i_B), \tag{2.55}$$

$$f_U(\mathbf{i}) = p(\mathcal{C}_T|\mathbf{i}) = p_A(\mathcal{C}_T|\mathbf{i}) \cdot p_B(\mathcal{C}_T|\mathbf{i}) = p(i_B|i_A) \cdot p(i_A|i_B) = \frac{p(\mathbf{i})^2}{p(i_A) \cdot p(i_B)}. \quad (2.56)$$

Here, $\mathcal{C}_T$ is a set of all intensities that represent correct match, while $p_A(\mathcal{C}_T|\mathbf{i})$ and $p_B(\mathcal{C}_T|\mathbf{i})$ are probabilities that intensity pair $\mathbf{i}$ represents correct match according to images A and B, respectively, which are actually conditional probabilities. Probability $p_B(\mathcal{C}_T|\mathbf{i})$ is equivalent to point similarity function $f_{PC}(\mathbf{i})$, see (2.41). The difference between similarity functions f_U and f_{PC} is therefore in conditional probability $p(i_B|i_A)$, which is expected to improve registration of non-dominant tissues. Although grouping intensities into classes is not used, measure S_U still incorporates some knowledge about joint distributions, adopted from segmentation based approach.

It cannot be expected that a certain intensity in one image corresponds to only one intensity in the other image, in the same way as it can happen for mean intensity values of tissue types (μ_A and μ_B) in case of a segmentation based measure. Each intensity in one image corresponds to several intensities in the other image. Estimated probabilities $p(\mathcal{C}_T|\mathbf{i})$ depend on the amount of noise, which can differ according to tissue type. To reduce the differences between similarities of correctly registered intensity pairs, a logarithmic function can be applied in the same way as in case of entropy based point similarity measure $S_H(\mathbf{i})$, to depend on uncertainty rather than probability. So we get a new point similarity measure S_{UH} with the following point similarity function:

$$f_{UH}(\mathbf{i}) = \log p(\mathcal{C}_T|\mathbf{i}) = \log \left(p\left(i_B|i_A\right) p\left(i_A|i_B\right)\right) = \log \frac{p(\mathbf{i})^2}{p(i_A) \cdot p(i_B)}. \quad (2.57)$$

The obtained point similarity measure S_{UH} is related to previously described measures S_H and S_{MI}, such that

$$f_{UH}(\mathbf{i}) = \log p(\mathbf{i}) + \log \frac{p(\mathbf{i})}{p(i_A) \cdot p(i_B)} = f_H(\mathbf{i}) + f_{MI}(\mathbf{i}). \quad (2.58)$$

2.5 Benefits of point similarity measures

Point similarity measures were designed for high-dimensional non-rigid registration and provide some advantages over the conventional similarity measures. The most important advantage is that they provide the two required properties: locality and multi-modality. In addition there are other advantages, which can contribute not only to high-dimensional registration, but also to low dimensional, e.g. rigid registration. We focus on advantages of point similarity measures in detail in the following subsections.

2.5.1 Locality and spatial deformation models

Here we focus on the phenomena that similarity measures may constrain the deformation and as such act as a spatial deformation model. An illustrative example is the work of Maintz [52], who performed non-rigid image registration without specific spatial deformation model. The only regularization comes from the similarity measures, which constrain the deformation because of the spatial voxel interdependencies used during the similarity measurement. Explicitly, the constraint is in the similarity measurement that uses large image regions, where region by itself cannot deform during the similarity measurement. The larger the regions are, the more constraint the deformation is. This can be explained using the Bro-Nielsen's approach [5], by which most spatial deformation models can be explained and implemented by convolution filtering, see Chapter 4.

Similarity of an image region or a whole image S_R can be computed from point similarities $S(\mathbf{x}) = S(\mathbf{x}, \mathbf{x})$ by averaging:

$$S_R = \overline{S(\mathbf{x})} = \frac{1}{N_R} \sum_{\mathbf{x} \in R} S(\mathbf{x}), \tag{2.59}$$

where R denotes image region for which similarity is being measured and N_R is number of estimated point similarities in this region. This relation can also be expressed as convolution filtering with some spatial filter G_R,

$$S_R(\mathbf{x}) = G_R * S(\mathbf{x}). \tag{2.60}$$

The larger the region is, the wider is the impulse response of the filter G_R and more global information is extracted from the point similarities. For example, when the region spreads over the whole images only global image properties, appropriate for global registration (e.g. rigid) are extracted, while detailed local information is blurred. Such extraction of more global knowledge from multiple more localized image features is characteristic for spatial regularization. When measuring region similarity this regularization is substantial and shall be considered when designing spatial deformation models.

As explained in Chapter 4., the majority of spatial deformation models can also be performed by convolution filtering [5]. For example, an incremental elastic registration can be performed iteratively as follows:

$$\mathbf{T}(\mathbf{x})^{t+1)} = \mathbf{T}(\mathbf{x})^{(t)} + G_E * \frac{\partial S(\mathbf{x})}{\partial \mathbf{T}(\mathbf{x})}, \tag{2.61}$$

where G_E denotes a filter with impulse response of the elastic media and t is the iteration number. When spatial deformation model is used in combination with the similarity of image region S_R, the Eq. (2.61) can be rewritten:

$$\mathbf{T}(\mathbf{x})^{(t+1)} = \mathbf{T}(\mathbf{x})^{(t)} + G_E * \frac{\partial G_R * S_P(\mathbf{x})}{\partial \mathbf{T}(\mathbf{x})} = \mathbf{T}(\mathbf{x})^{(t)} + G_E * G_R * \frac{\partial S_P(\mathbf{x})}{\partial \mathbf{T}(\mathbf{x})}. \tag{2.62}$$

Similar results could also be obtained for other spatial deformation models, e.g. viscous fluid or elastic. The regularization is therefore duplicated, which means that the final effect does not directly follow the spatial deformation model, see Fig. 2.4. Point similarity measures solve this problem. They push the limits of the locality into extreme and, consequently, similarity of one point does not presume any spatial relation with neighboring image points. Therefore, the transformations are not regularized by G_R, such that regularization remains solely in the domain of spatial deformation model, which gains a full control over the transformation properties.

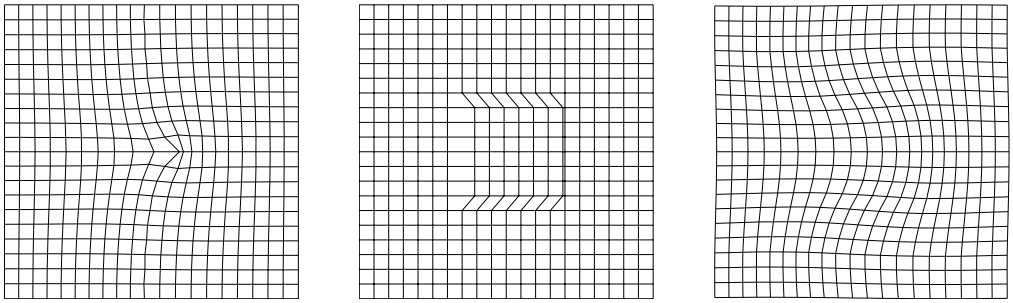


Figure 2.4: *An example of convolution filters used for regularizing non-rigid registration: elastic filter G_E (left), filter G_R that corresponds to region averaging (middle), and their convolution $G_E * G_R$ (right).*

Note that spatial regularization is necessary, otherwise matching of individual image points is ill-posed if they are matched independently [24, 52]. Because point similarity measures do not provide regularization implicitly, it must be provided by a suitable spatial deformation model. However, the final spatial deformation properties exactly follow properties defined by this model, as there is no regularization provided by the point similarity measures. This is not the case when using similarity measures with larger image regions.

2.5.2 Interpolation artifacts

Image registration methods search for such a transformation that maximizes the image similarity. Thus, it is important to study how the similarity changes with respect to the transformation. A common problem are the so-called interpolation artifacts, i.e. disproportional change of similarity with respect to image transformation, which may result in local similarity extrema that rules out the subvoxel accuracy and deteriorate the registration. In case of non-rigid image registration interpolation artifacts cannot be reduced using approaches proposed for rigid or low dimensional registration [63, 47], because they can only be applied when

comparing large image regions. Here we analyze interpolation artifacts of point similarity measures and compare them with interpolation artifacts of conventional similarity measures.

Let us suppose we have two images, A and B, and transform image B with transformation $\mathbf{T}$. The transformation moves each point $B(\mathbf{x})$ from its original position $\mathbf{x}$ for some displacement $\mathbf{T}(\mathbf{x})$ to a new position $\mathbf{x}+\mathbf{T}(\mathbf{x})$, where it gets matched with a point $A(\mathbf{x}+\mathbf{T}(\mathbf{x}))$. Let us observe a point similarity $S(v_B)$ at voxel point v_B in image B, where

$$S(v_B) = S(A(\mathbf{x}+\mathbf{T}(\mathbf{x})), B(\mathbf{x})). \tag{2.63}$$

This point similarity can change due to two reasons. The first one is the change of point pair, and the second one is a possible change of the joint intensity distribution, which changes the point similarity function $f(\mathbf{i})$. Both of them could potentially cause interpolation artifacts.

First we focus on the changes of point similarity due to the changes of the point pair and assume that the point similarity function $f(\mathbf{i})$ does not change. Thus, point similarities $S(v_B)$ change only because points in image B are compared with different points in image A. However, due to the discrete nature of the images and due to the image transformation, voxel points in image B do not match exactly with voxel points in image A, and measuring of point similarities requires interpolation. In case of mutual information there are two interpolation methods commonly used: interpolation of intensity and partial volume interpolation. Interpolation of intensity can also be employed in case of point similarities such that:

$$S(v_B) = f(i_A(\mathbf{x}+\mathbf{T}(\mathbf{x})), i_B(v_B)), \tag{2.64}$$

where the unknown intensity $i_A(\mathbf{x}+\mathbf{T}(\mathbf{x}))$ is interpolated from intensities of neighboring voxels:

$$i_A(\mathbf{x}+\mathbf{T}(\mathbf{x})) = i_A(v_B) = \sum_{v_A} \omega(v_A, v_B) i_A(v_A). \tag{2.65}$$

Here ω stands for weights used for the interpolation. However, the interpolation of intensity assumes a linear intensity dependence, which may not necessarily comply with the intensity dependence estimated from the images, and can cause interpolation artifacts, as shown in Fig. 2.5. To avoid the interpolation artifacts, we propose to interpolate point similarity instead of intensity:

$$S(v_B) = \sum_{v_A} \omega(v_A, v_B) S(v_A, v_B), \tag{2.66}$$

where the weights ω are the same as in the case of interpolating intensity and $S(v_A, v_B)$ is a point similarity between a point at voxel v_A in image A and a point at voxel v_B in image B. This approach results in a linear relationship between

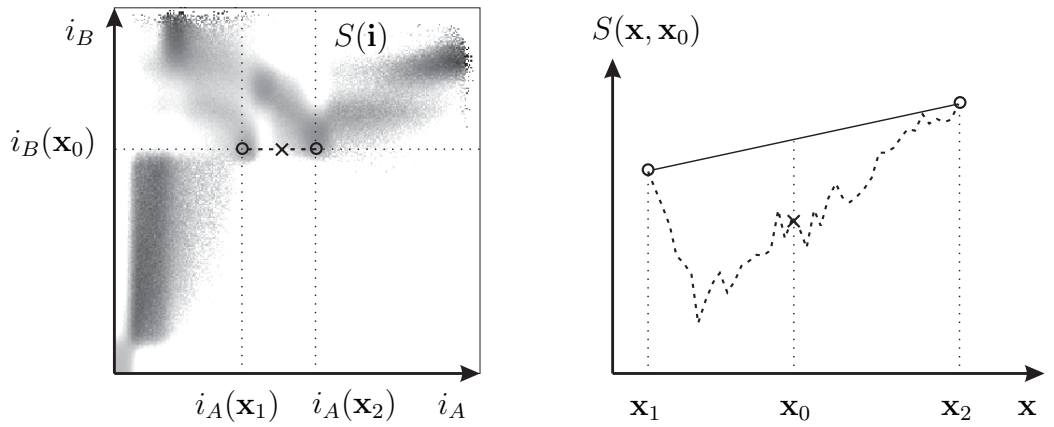


Figure 2.5: *Illustration of measuring point similarity with interpolation. Similarity between some voxel point $B(\mathbf{x}_0)$ and corresponding point $A(\mathbf{x}_0)$ requires interpolation. In case of interpolation of intensity, an intensity $i_A(\mathbf{x}_0)$ is interpolated from intensities of neighboring points $i_A(\mathbf{x}_1)$ and $i_A(\mathbf{x}_2)$, and the point similarity is $S(\mathbf{x}_0, \mathbf{x}_0) = S(i_A(\mathbf{x}_0), i_B(\mathbf{x}_0))$ (cross mark). Because the interpolated intensity does not comply with the complex intensity dependence estimated from the images, the similarity does not have a correct meaning, which introduces interpolation artifacts (dashed line). The problem can be solved by interpolating point similarity instead of intensity, such that $S(\mathbf{x}_0, \mathbf{x}_0)$ is interpolated from point similarities $S(\mathbf{x}_1, \mathbf{x}_0)$ and $S(\mathbf{x}_2, \mathbf{x}_0)$. In this case interpolation of intensity is not required and interpolation artifacts do not appear (solid line).*

the point similarity and point displacement in a range of one image voxel, thereby avoiding the interpolation artifacts, see Fig. 2.5. The difference between results obtained by using different interpolation methods is illustrated in Fig. 2.6.

When interpolation of point similarities is used to measure similarity of the whole images by averaging point similarities, it is equivalent to the partial volume interpolation, which is often used for global statistical measures, e.g. mutual information. In order to prove the equivalence let us first mathematically express the partial volume interpolation:

$$p(\mathbf{j}) = p(j_A, j_B) = \frac{1}{N} \sum_{v_B} \sum_{v_A} \omega(v_A, v_B) \delta(i_A(v_A) - j_A) \delta(i_B(v_B) - j_B), \qquad (2.67)$$

where N is the number of all image voxels, δ is a Dirac's delta function and $\omega(v_A, v_B)$ is a weight used for interpolating point at voxel v_B with voxel points of image A. The same weights ω are used for interpolating point similarities. Thus the global similarity S_R, computed by averaging point similarities $S(v_B)$ can be

expressed as follows:

$$\begin{aligned} S_R &= \frac{1}{N}\sum_{v_B} S(v_B) \\ &= \frac{1}{N}\sum_{v_B}\sum_{v_A} \omega(v_A, v_B) f(i_A(v_A), i_B(v_B)) \\ &= \frac{1}{N}\sum_{v_B}\sum_{v_A} \omega(v_A, v_B) \sum_{\mathbf{j}=[j_A,j_B]} f(\mathbf{j})\delta(i_A(v_A) - j_A)\delta(i_B(v_B) - j_B) \\ &= \sum_{\mathbf{j}=[j_A,j_B]} f(\mathbf{j}) \frac{1}{N}\sum_{v_B}\sum_{v_A} \omega(v_A, v_B)\delta(i_A(v_A) - j_A)\delta(i_B(v_B) - j_B) \\ &= \sum_{\mathbf{j}=[j_A,j_B]} f(\mathbf{j})p(\mathbf{j}). \end{aligned} \tag{2.68}$$

In the case that point similarity function $f(\mathbf{i})$ is also computed from the same joint distribution $p(\mathbf{i})$, then according to the Eq. 2.14 the global similarity S_R equals the global similarity S_G, obtained by corresponding global similarity measure by using the same joint distribution $p(\mathbf{i})$.

$$S_R = \sum_{\mathbf{i}} f(\mathbf{i})p(\mathbf{i}) = S_G. \tag{2.69}$$

Note that $p(\mathbf{i})$ was obtained by partial volume interpolation, such that Eq. 2.69 proves our statement about the equivalence of partial volume interpolation and interpolation of point similarity.

The second possible source of changes in point similarity is the change of the intensity distributions. In general, the transformation $\mathbf{T}$ changes the marginal and joint intensity distributions $p(i_A)$, $p(i_B)$ and $p(\mathbf{i})$. If the point similarity function $f(\mathbf{i})$ is recomputed using the updated distributions, then it changes as well and so do the point similarities. The relation between the transformation and the change of intensity distributions is complex, nonlinear and depends on the information of the whole images. Furthermore, the relation between the intensity distributions and the corresponding point similarity function is usually not linear either, see Eq. (2.38). The nonlinear relation reflects in interpolation artifacts, which cannot be avoided when using conventional multi-modality similarity measures, e.g. mutual information. Nevertheless, point similarity function $f(\mathbf{i})$ is in all cases, not regarding the transformation $\mathbf{T}$, an approximation of the same intensity dependence, i.e. the intensity dependence of correctly registered images. The only difference between the obtained point similarity functions is in the quality of the estimation, which depends on the level of global image mismatch. However, the change of point similarity caused by reestimation of point similarity function does not provide reliable information about the quality of image match, because the changes of point similarity function do not originate only in the quality of

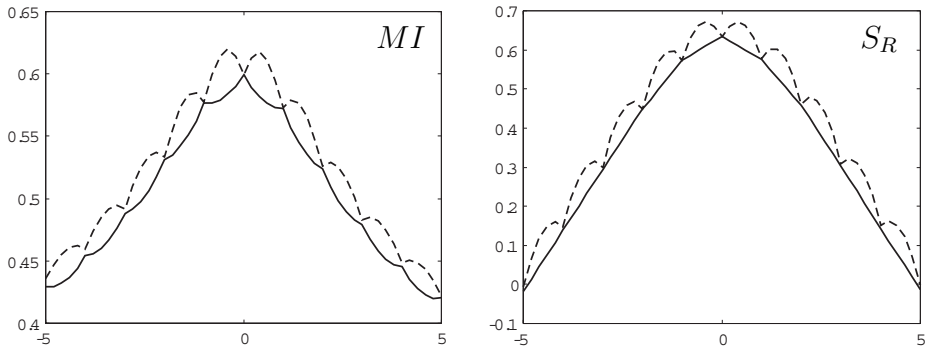


Figure 2.6: *An example of mutual information MI (left) and point based similarity $S_R = \overline{S_{MI}}$ (right) with respect to image translation, for the two different interpolation methods. The dashed lines denote interpolation of intensity, while the solid lines denote partial volume interpolation for MI (left) and interpolation of similarity for S_R (right). The difference between MI and S_R is due to using the same point similarity function $f_{MI}(\mathbf{i})$ for all the estimations of point similarity, not regarding the image translation. Keeping the point similarity function fixed therefore enables point similarity measures to completely avoid the interpolation artifacts. This cannot be achieved when using MI.*

image match, but also in the artifacts caused by image processing algorithms, e.g. interpolation used to obtain the intensity distributions. Consequently, in order to avoid interpolation artifacts we propose to estimate point similarities, without recomputing the similarity function. In support of this approach we show that registration based on point similarity measures always tends towards the transformation that would be obtained when using point similarity function estimated from the registered images, see Fig. 2.8. When the point similarity function is estimated at better match, the similarity better distinguishes between correct matches and mismatches, while the positions of maxima that correspond to different tissue types in the point similarity function do not change. This is illustrated in the experiment performed using simulated Brainweb images [42] and mutual information based point similarity measure S_{MI} in Fig. 2.7. Note that the quality of image match used for estimation of point similarity function affects the sensitivity of point similarity, such that the sensitivity decreases with increasing image mismatch. Therefore, to avoid the interpolation artifacts and still achieve good sensitivity, we propose to use the same point similarity function for all the measurements of similarity that need to be compared, and recompute point similarity function only once per registration step or registration iteration.

To summarize, in order to avoid the interpolation artifacts we recommend to keep the point similarity function $f(\mathbf{i})$ fixed and use interpolation of similarity instead of interpolation of intensity. This avoids all the interpolation artifacts of point similarity measures, as illustrated in Figure 2.6. Conventional statisti-

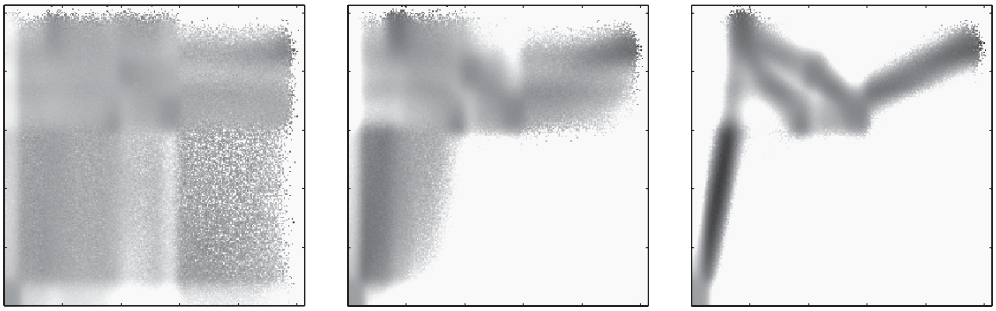


Figure 2.7: *Point similarity functions for simulated MRI-T1 and MRI-PD images of the head, at different levels of image mismatch: 10 mm displacement (left), 2 mm displacement (middle), and registered images (right). Darker color represents higher similarity. Note that the positions of maxima that correspond to different tissue types in point similarity function do not change.*

cal measures, e.g. mutual information, cannot avoid interpolation artifacts in a similar way, because they always implicitly reestimate the intensity dependence between the images. This suggests that point similarity measures can contribute not only to high dimensional registration, which they were designed for, but also to low-dimensional registration, which also faces with the problem of interpolation artifacts.

2.6 Comparison of point similarity measures

So far we have described different methods to estimate point similarity functions and show some advantages of the point similarity approach. However, one important question remains: *Which similarity measure to use?* In this section we aim to answer this question. First we present a general strategy for choosing a similarity measure in Subsection 2.6.1, then we compare the similarity measures with respect to their similarity functions in Subsection 2.6.2 and finally, we compare them as an integral part of a high-dimensional non-rigid registration system in Subsection 2.6.3.

2.6.1 Selection of point similarity measures

An important question is which similarity measure to use for solving a given registration problem. It is well known that some similarity measures, in particular mutual information measures, are so general that they can be used for almost any kind of images. However, better results may be obtained by using appropriately more constrained measures. Estimation of a parameter that is known in advance is not reasonable, as the result may be biased due to its imperfect estimation.

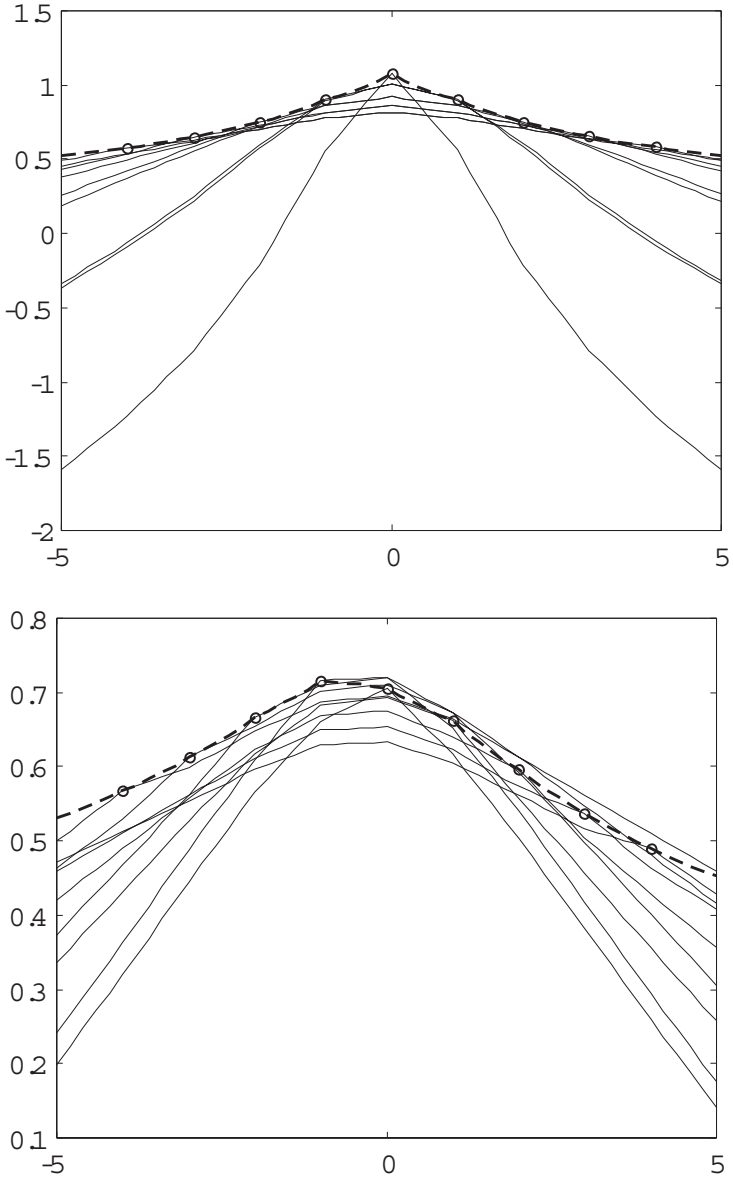


Figure 2.8: *Mutual information (MI) (dashed line) and similarities obtained using the mutual information based point similarity measure based on different estimations of point similarity function $S(\mathbf{i})$ (solid lines), with respect to image displacement. Similarity is measured between MRI-PD and MRI-T1 data, using simulated images (top) and real images of the head (bottom). Point similarity functions $f_{MI}(\mathbf{i})$ were estimated at different image displacements. At this displacement the point based similarity equals the mutual information (marked with circles). All point based similarities reach the maximum at displacement 0, i.e. where images are correctly registered. The experiment based on real images (bottom) shows that point similarity measures can succeed even in the cases when MI fails.*

Moreover, estimation of additional parameters requires additional information, or the uncertainty of results increases. Undoubtedly, the best results can be expected when the model fits well to the real situation.

Selection of a similarity measure should follow the same principle. Specifically, a similarity measure should be chosen according to the characteristics of the images. We have already classified similarity measures into mono-modality and multi-modality. However, additional classifications are possible [65]. Mono-modality similarity measures could be further classified according to model assumptions into $i_B = i_A$, $i_B = k \cdot i_A$ or $i_B = k \cdot i_A + n$. Similarly, multi-modality similarity measures are sometimes classified into measures with functional intensity dependencies and those with statistical intensity dependencies [65]. Furthermore, multi-modality similarity measures could be classified according to the model assumptions into those that model intensity classes, and those where no intensity classes are modeled.

Let us illustrate the selection of similarity measures with some examples. Firstly, non-rigid registration of images taken in a single time series can often presume mono-modality intensity dependency $i_B = i_A$, and thus it is expected that best results can be obtained when using similarity measures MAD, MSD or their derivatives. Secondly, for registration of 3D anatomical images, in the case that intensity classes can be modeled, good results are expected by using point similarity measure S_S (2.43). Thirdly, when intensity classes cannot be modeled, measures S_{MI} (2.38), S_U (2.56) or S_{UH} (2.57) are preferred. Finally, images that include functional information may use S_{MI}, S_U or S_{UH} measures as well.

2.6.2 Comparison of similarity functions

A reasonable approach to compare global similarity measures is to observe similarities according to the applied image transformation, e.g. displacement of whole image, and evaluate some properties of such similarity function, e.g. number of local extrema, smoothness, position of global extremum, capture range, etc [62]. However, comparison of point similarity measures turns out to be more problematic as any transformation of such an extremely small image region makes a drastic change in region overlap. This makes such comparison impossible. However, point similarity measures can be compared in the intensity domain instead of the transformation domain. Thus, in this subsection we compare point similarity measures according to their similarity functions.

Similarity functions are estimates of intensity dependence between the images, when they are registered. Ideally, they should have high and equal values (similarity) for all intensity pairs that represent correct matching (true classes) and furthermore, they should have low and equal values for all other intensity pairs (including false classes). There are two issues that are especially important for

the non-rigid registration. The first one is the ratio between similarity for correct and incorrect match, which should be low to enable good distinction between correct and incorrect matching and the second one is the ratio between similarity for different correctly matched intensity pairs, which should be close to 1 in order to treat all tissue types equally, such that all of them could get registered equally well, and prevent domination of some tissue types over the others.

In this subsection we analyze and compare point similarity measures according to their similarity functions. We analyze and compare all previously described measures: S_P, S_H, S_{PC}, S_{HC}, S_U, S_{UH}, S_{MI} and S_S. For each of them we compute the similarity functions for two cases. First, for correctly matched images, where joint intensity distribution consists only of true classes, which represent correctly matched tissues, and second for mismatched images, which were displaced for 10mm, such that the additional false intensity classes that represent incorrect match appear as well. All these experiments are performed using Brainweb simulated images of human head, specifically, normal high resolution MRI-T1 and MRI-PD images. The results are shown in Figures 2.9 to 2.16. Which intensity pair belong to which tissue type pair is shown in Fig. 2.3.

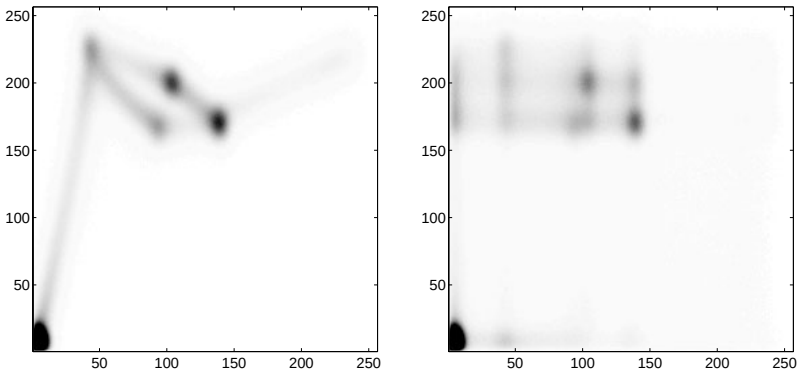
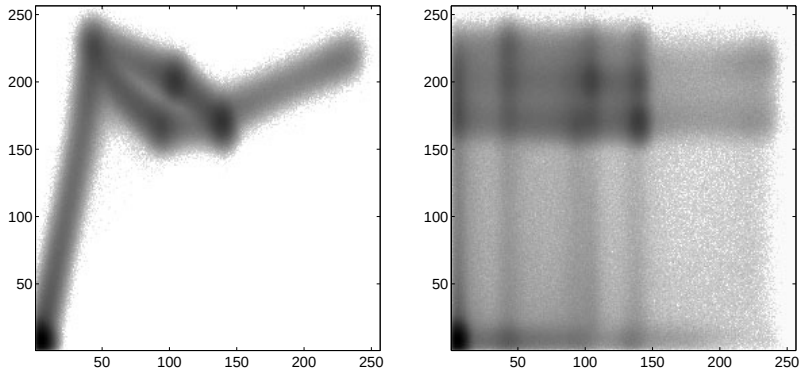


Figure 2.9: *Point similarity function $f_P(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

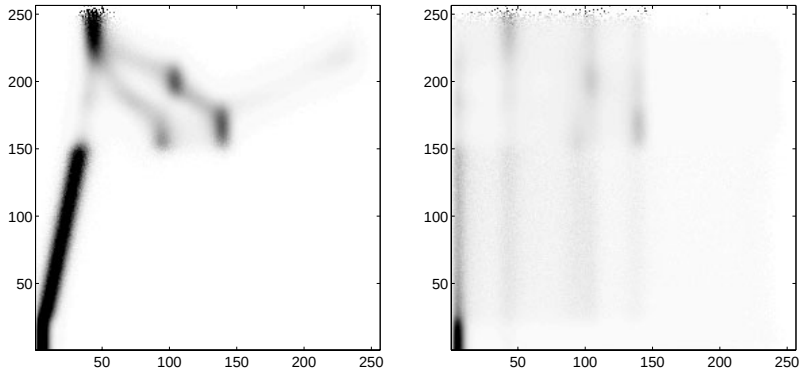
Point similarity function $f_P(\mathbf{i})$, see Fig. 2.9, of point similarity measure S_P, is directly an estimate of a current joint intensity distribution. The similarity for certain intensity pair is therefore linearly related with the amount of tissue that is represented with these intensities. Consequently, the similarity of more frequent tissues is higher than the similarity of infrequent tissues, such that the quality of registration of the later ones are expected to be deteriorated due to the dominance of the first ones. In addition, this point similarity function poorly distinguishes between intensity pairs that represent true and those that represent false matching.



(a) (b)

Figure 2.10: *Point similarity function $f_H(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

Point similarity measure S_H is related to S_P. Its point similarity function $f_H(\mathbf{i})$, see Fig. 2.10, is only a non-linearly scaled version of $f_P(\mathbf{i})$, using a log function. Consequently, its ability to distinguish between correct and incorrect match is poor as well. However, the log function makes the similarity for different tissues less dependent on their amount, and so, the improvement of registration can be expected.



(a) (b)

Figure 2.11: *Point similarity function $f_{PC}(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

Point similarity function $f_{PC}(\mathbf{i})$, see Fig. 2.11, of measure S_{PC}, tends to make the similarities for different tissues more similar by normalizing the joint intensity distribution with the marginal intensity distribution of image B. This can

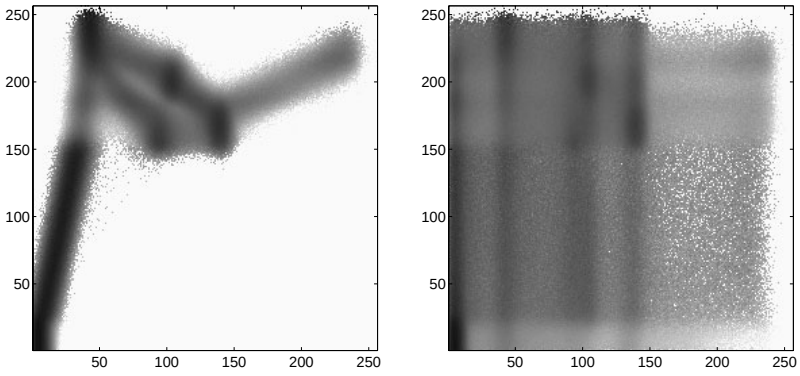


Figure 2.12: *Point similarity function $f_{HC}(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

be justified by the fact that the majority of registration techniques estimate the external forces as a gradient of similarity of certain point in one image (B) according to its displacement with respect to the other image (A). However, the difficulties appear when certain intensity of image B could correspond to various intensities of image A, e.g. intensities that represent different tissues. Such situation is when the point similarity function is obtained from mismatched images. In this case, the similarities of different correctly matched tissues may differ considerably, and furthermore, the discrimination between correct and incorrect matches is relatively bad.

Point similarity function $f_{HC}(\mathbf{i})$, see Fig. 2.12, of measure S_{HC} is a logarithmicaly scaled version of the similarity function $f_{PC}(\mathbf{i})$. This makes similarities of different correctly matched tissues more similar, but however, this also deteriorates discrimination between correct and incorrect match.

Point similarity function $S_U(\mathbf{i})$, see Fig. 2.13, for measure S_U, could be understood as an improvement of $f_{PC}(\mathbf{i})$, as it uses normalization with marginal intensity distributions. The difference is that in this case normalization is performed twice, i.e. with both marginal distributions. This considerably improves the discrimination between correct and incorrect matching. However, it also shows larger differences between similarities of different correctly matched tissues.

Point similarity function $f_{UH}(\mathbf{i})$, see Fig. 2.14, for measure S_{UH}, is a logarithmicaly scaled version of the similarity function $f_U(\mathbf{i})$. It retains a good discrimination between correct and incorrect matches, but makes similarities of different correctly matched tissues more equal. This measure is therefore one of the best similarity measures in both of the observed aspects.

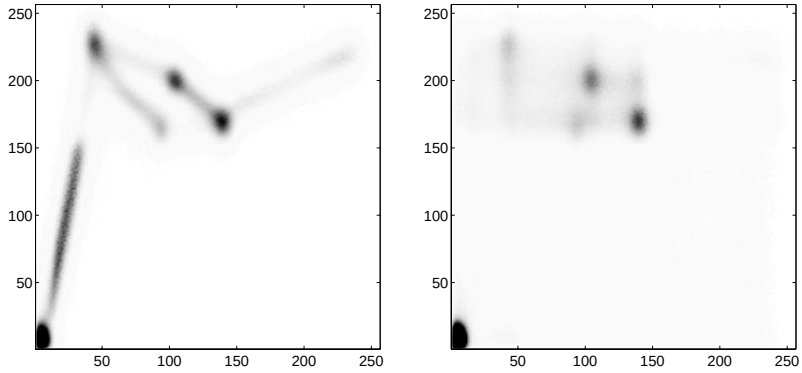


Figure 2.13: *Point similarity function $f_U(\mathbf{i})$, obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

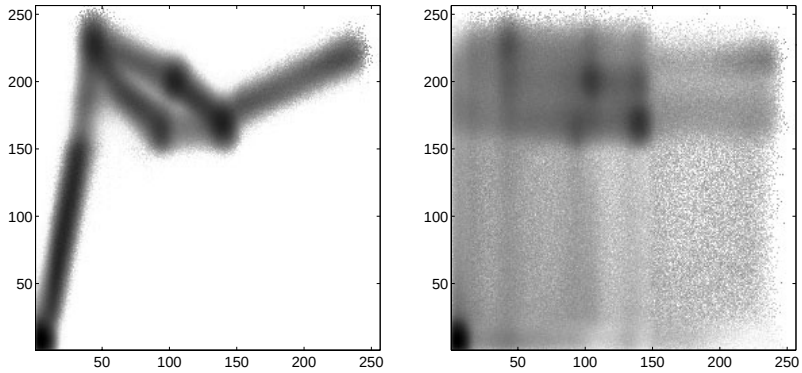


Figure 2.14: *Point similarity function $f_{UH}(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

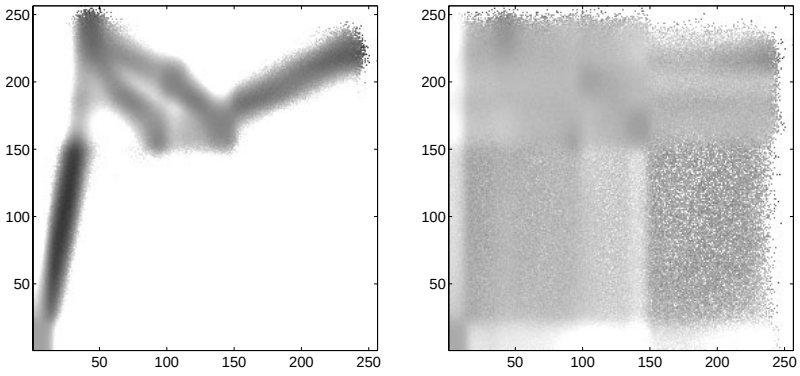


Figure 2.15: *Point similarity function $f_{MI}(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

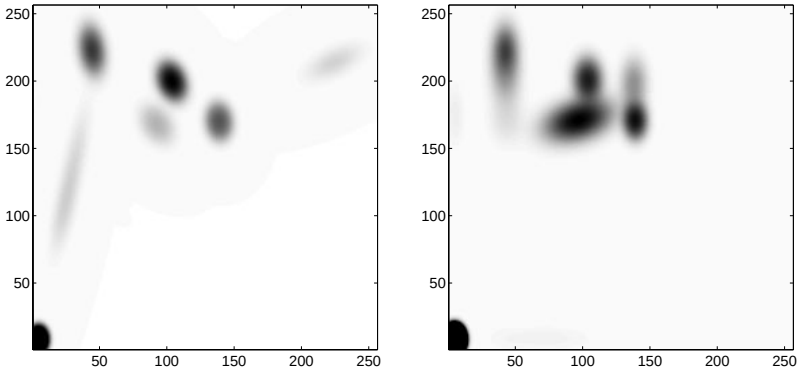


Figure 2.16: *Point similarity function $f_S(\mathbf{i})$ obtained from matched images (a), and mismatched images (b), for MRI-T1 and MRI-PD images of head. Darker color represents higher similarity.*

Point similarity function $f_{MI}(\mathbf{i})$, see Fig. 2.15, for measure S_{MI} is derived from mutual information. According to this, events with lower probability provide higher information. In the case of point similarity measures this means that tissue types with lower probability dominate over more frequent tissues. In this manner $f_{MI}(\mathbf{i})$ is exactly the opposite to all other point similarity functions, where more frequent tissues obtain higher similarity. However, $f_{MI}(\mathbf{i})$ is good in booth observed aspects, as it also provide relatively good discrimination between correct and incorrect matching.

Point similarity function $f_S(\mathbf{i})$, see Fig. 2.16, for measure S_S, considerably

differs from other point similarity functions as it does not directly model partial volume voxels. They are modeled only if they form a peak in joint intensity distribution. However, the strength of $f_S(\mathbf{i})$ is in good discrimination between correct and incorrect match (true and false classes) as well as in no discrimination between the true classes. However, this similarity function can be used only in the cases when joint distribution can be modeled as a sum of intensity classes. In our case the classes are approximated with Gaussian functions, although this is in nature often not absolutely correct. This makes it sensitive to intensity inhomogeneity, and furthermore, the problems may appear when classes are very close to each other. The later situation is evident in the example in Fig. 2.14 (b), where the true class for surrounding tissues S (see Fig. 2.3) and a false class for GM/WM are modeled as a single class, i.e. with a single Gaussian function, and thus, they cannot be distinguished well.

To summarize, measures S_P, S_{PC} and S_U highly depend on the amount of different tissue types. In our case background dominates over the tissues and thus registration of tissue intensities is much slower than registration of background, which can result in relatively bad registration results. A log function, which is used for measures S_H, S_{MI}, S_{HC} and S_{UH}, makes similarities for different correctly registered tissues less different. However, similarities for incorrectly registered parts become more similar to those for correct match as well. This becomes problematic especially in case of measure S_{HC}. Similarity functions for S_H, S_{MI} and S_{UH} are related (2.58). S_H is directly related to joint intensity distribution $p(\mathbf{i})$ and produces high similarity for intensity pairs with high probability $p(\mathbf{i})$, e.g. for true classes. On the other hand measure S_{MI} provides good discrimination between true and false classes, while similarity is higher for less frequent tissues and relatively low for more frequent ones, e.g. for background. This is disadvantageous when high tissue misalignment exists, which is common for initial registration steps. Measure S_{UH} improves that, and equalizes the similarities of true classes, while discrimination between true and false classes remains good. Similarity function for measure S_S differs from other measures, as it models the joint distribution as a sum of intensity classes, while other intensity pairs, e.g. partial volume voxels, are not modeled. In comparison to the other point similarity measures it better distinguishes between true and false classes.

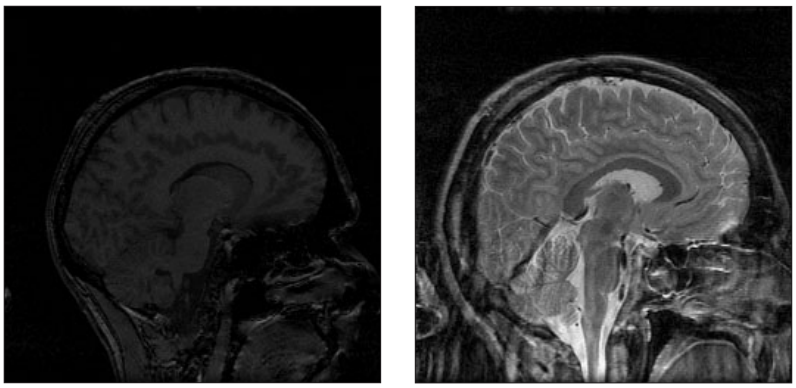
Overall, according to the distinction between correct and incorrect matches, and likeness between similarities for correctly matched tissues, the most promising are measures S_S for images where intensity distribution can be modeled with classes and S_{UH} for other cases.

2.6.3 Comparison of registration results

The most meaningful way to compare point similarity measures is to compare registration results obtained using these measures. To perform such a compari-

son we used our system for multi-modality non-rigid image registration, which is described in Chapter 6. To show performance of point similarity measures under different registration configurations, we have compared them using two force estimation methods: a standard approach, which uses forward forces only, and symmetric approach, which uses forward and reverse forces and is described in Chapter 3. In the experiments we used four resolution levels, each consisting of 10 registration iterations and spatial deformation model with Gaussian filters with standard deviations $\sigma_{\mathbf{G}1} = \sigma_{\mathbf{G}2} = 3$.

In the experiment we used simulated and real images. First we have used MRI T1 and PD Brainweb simulated images [42] with $1 \times 1 \times 1$ mm voxel size. The tests were performed using images of a whole head and images of brain only, both of them in three different image qualities: normal images, 9% noisy images and 40% intensity inhomogeneous images. In total 6 simulated image pairs were used. The same experiment was performed using real MRI images. For that purpose we used MRI T1 image as target with voxel size $0.86 \times 0.86 \times 0.99$ mm, while the MRI PD source image had $0.98 \times 0.98 \times 1.1$ mm voxel size. The images were previously corrected for intensity inhomogeneity using the information minimization approach [48]. The comparison based on real images is more demanding, because the quality of real images is worse than the quality of simulated images. Namely, target image (MRI-T1) had poor intensity resolution, while source image (MRI-PD) included large image artifacts (intensity inhomogeneity and wrap arround artifacts), see Figure 2.17.



(a) (b)

Figure 2.17: *Real images used for evaluation. The target image MRI-T1 (a) has poor intensity resolution, while source image MRI-PD (b) includes large image artifacts, i.e. intensity inhomogeneity and wrap around artifacts.*

Our comparison of multi-modality point similarity measures is based on recovering a synthetic deformation, see Chapter 5., Section 5.3.1. We used a synthetic deformation generated as a sum of Gaussian functions. Specifically, six functions with standard deviation ranging between 15 and 60 mm were used, resulting in

Table 2.1: *Comparison results for simulated images of a whole head. All results are in millimeters and represent residual errors after the registration.*

Forward force estimation:

measure	normal		9% noise		40% shading	
	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}
S_P	5.35	14.21	5.08	14.51	6.15	15.32
S_H	0.89	5.39	0.99	5.77	22.70	44,09
S_{MI}	1.09	6.83	1.23	7.29	3.00	10.01
S_{PC}	10.20	22.04	15.69	26.21	22.63	43.64
S_{HC}	0.86	5.35	1.03	6.14	23.02	44.56
S_S	0.68	2.44	0.73	3.09	3.72	10.71
S_U	2.78	10.42	3.17	11.41	5.97	15.01
S_{UH}	0.83	4.90	1.09	6.36	2.68	9.18

Symmetric forces:

measure	normal		9% noise		40% shading	
	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}
S_P	3.67	12.32	3.48	12.18	4.65	13.54
S_H	0.80	4.50	0.87	4.93	7.47	24.13
S_{MI}	0.83	5.04	0.99	6.05	2.39	8.09
S_{PC}	1.05	6.87	1.47	8.35	3.11	10.07
S_{HC}	0.71	4.23	0.91	5.43	15.69	37.85
S_S	0.51	1.91	0.55	2.08	4.02	11.12
S_U	2.00	9.22	2.00	9.31	4.55	13.21
S_{UH}	0.72	4.08	0.79	4.59	2.59	9.25

initial displacement error $e_{rms} = 6.90$ mm for simulated Brainweb images, and $e_{rms} = 14.15$ mm for real images. The difference between real and simulated images arose due to different position and orientation of head. We cannot argue that Gaussian deformation can be used to model actual deformations in clinical multi-modal applications. However, due to functional independence of force estimation stage and deformation model used, the results obtained that way are appropriate for comparing force estimation methods and their fundamental part - similarity measures. While the spatial deformation model may favor some type of deformation, ordering of different external force estimation approaches is still preserved.

The results of registering simulated images are tabulated in Table 2.1 and Table 2.2. For all compared similarity measures sensitivity to noise is low while sensitivity to intensity inhomogeneity (shading) is relatively high. The most sensitive to intensity inhomogeneity was registration of whole head images with

Table 2.2: *Comparison results for simulated images of brain only. All results are in millimeters and represent residual errors after the registration.*

Forward force estimation:

measure	normal		9% noise		40% shading	
	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}
S_P	5.20	13.24	5.72	14.78	7.03	16.59
S_H	0.73	3.99	0.99	5.88	2.14	8.95
S_{MI}	1.65	9.17	1.79	9.32	2.42	10.18
S_{PC}	0.72	3.48	0.95	3.50	3.01	9.98
S_{HC}	1.29	7.69	1.54	8.50	2.24	9.61
S_S	0.55	2.19	0.77	2.58	3.96	11.82
S_U	1.57	4.25	1.98	4.49	8.04	17.22
S_{UH}	0.56	2.56	0.68	3.23	1.54	5.96

Symmetric forces:

measure	normal		9% noise		40% shading	
	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}
S_P	3.35	12.18	3.36	12.21	3.42	12.29
S_H	0.50	2.50	0.65	3.36	2.09	9.45
S_{MI}	0.95	6.07	1.11	6.89	1.92	9.48
S_{PC}	0.98	6.14	1.10	6.39	2.28	9.35
S_{HC}	0.65	3.40	0.79	4.33	1.91	8.94
S_S	0.48	1.98	0.59	2.19	3.22	10.14
S_U	0.63	2.49	0.84	2.77	7.72	18.65
S_{UH}	0.43	2.10	0.49	2.23	1.77	7.98

Table 2.3: *Comparison results of multi-modal point similarity measures using simulated deformation and real MRI-PD/T1 images of human head for forward force estimation ($\mathbf{F} = \mathbf{F}_F$) and for symmetric forces ($\mathbf{F} = \mathbf{F}_F - \mathbf{F}_R$). All results are in millimeters.*

measure	$\mathbf{F} = \mathbf{F}_F$		$\mathbf{F} = \mathbf{F}_F - \mathbf{F}_R$	
	e_{RMS}	e_{max}	e_{RMS}	e_{max}
S_P	10.17	22.95	7.11	21.17
S_H	2.67	13.80	2.07	10.91
S_{MI}	2.17	12.23	2.11	10.99
S_{PC}	11.41	25.68	2.36	11.54
S_{HC}	2.67	13.79	2.08	9.87
S_S	1.56	6.56	1.66	7.33
S_U	8.04	20.21	5.44	18.80
S_{UH}	2.32	12.50	1.83	10.22

forward force estimation, where registration was successful only when using measures S_{MI}, S_S and S_{UH}. Note that the level of intensity inhomogeneity was higher than is expected for real data. The problem appears due to intensities of tissues surrounding brain, which are similar to intensities of some brain tissue in the source image, and some other tissue in the target image. In such a situation registration can deteriorate, especially when it is based on measures S_{PC}, S_U and S_P, and in more extreme cases also measures S_H and S_{HC}. Shading has smaller influence on e_{RMS} for images of brain, as they do not include surrounding tissues, and consequently, class overlap at intensity shading is lower than in the case of images of whole head, where surrounding tissues form additional intensity classes. In general, when images are not subject to intensity inhomogeneity, the best results are mostly obtained using measure S_S, in other cases the best results are mostly obtained by measure S_{UH}. When images include only intensity classes that are easy to distinguish, e.g. in case of registering brain images, almost every measure gives good results and differences between them are less apparent. The results also show that symmetric forces in general give better results than forward forces.

Results obtained for registering real images, which are tabulated in Table 2.3, are not as good as results obtained for the simulated images. There are two possible reasons. First, quality of real images is worse that quality of simulated images, and second, worse results may be caused by higher initial misalignment. However, comparison of results obtained from real images suggest similar conclusions as we made for simulated images. Segmentation based measure S_S performed best, measures S_{MI} and S_{UH} were also good, especially for consistent force esti-

mation. On the other hand measures S_P, S_U and eventually S_{PC} are shown to be less appropriate for such registration tasks.

2.7 Conclusion

In this chapter we described a novel approach for measuring similarity of extremely small image regions, which we named point similarity measures. We presented several point similarity measures, showed their advantages and discussed the obtained results. The most important advantage of the proposed approach is decoupling of spatial regularization from similarity measurement, which allows better control over the spatial deformation properties by the spatial deformation model. The next important advantage is the ability to avoid interpolation artifacts, which improves the convergence of registration. Furthermore, the computational cost of point similarity measures is low, which in practice enables higher dimensional and more precise registration. The comparison results of point similarity measures show that the best measures are the segmentation based point similarity measure S_S and point similarity measure S_{UH}. They gave the best results according to the point similarity function criteria and, furthermore, they performed best with respect to the obtained registration results. Due to the described advantages of point similarity measures and promising registration results obtained in the experiments, we found these measures suitable for multi-modality non-rigid registration. We anticipate that the quality of multi-modality registration based on point similarity measures can be compared to the quality of mono-modality registration.

3. Symmetric Image Registration

A quality of image match is usually estimated by measuring image similarity. However, similarity measures cannot assess transformations that do not change the appearance of the deformed image. In the case of non-rigid registration this causes differences between results obtained by registering images in different registration directions. This asymmetric relation leads to registration inconsistency and reduces the quality of registration. In this chapter we present a registration approach, which tends to improve the registration by establishing a symmetric image interdependence. In order to gather more information about the image transformation it measures the similarity in both registration directions. The presented solution is based on the interaction between the images involved in the registration process. Images interact with forces, which according to the Newton's action-reaction law form a symmetric relationship. These forces may transform both of the images, although in our implementation one of the images remains fixed. The experiments performed to demonstrate the advantages of the symmetric registration approach involve registration of simple objects, recovering synthetic deformation and interpatient registration of real images of the head. The results show that the symmetric approach improves the registration consistency and the registration correctness.

3.1 Introduction

The aim of image registration is to find a transformation that puts two imaged anatomies into a spatial correspondence. Each anatomical point in the first anatomy is expected to have exactly one homologous point in the second anatomy. It is intuitively expected that more correct registration reflects in higher image similarity, which is also a common assumption of image registration procedures. However, this assumption may not always be valid, as similarity measures compare images by their appearance and cannot asses transformations that do not change them in that aspect.

Let us suppose that the overall transformation consists of several transformation components. Such components can correspond to changes of individual transformation parameters. The problem arises when some of the components

do not cause any change in the image appearance and therefore they cannot be assessed by measuring image similarity. By increasing the dimensionality of the transformation the portion of transformation components that cannot be assessed by measuring similarity increases. This induces difficulties in non-rigid registration and its validation [89], see Chapter 5.

One of the properties of the similarity measurement is asymmetry. The results of measuring image similarity in different registration directions, i.e. by transforming the other image to obtain the same match, may depend on different transformation components. This is illustrated in Figure 3.1. Let us first assume that image A is being registered to image B. The transformation $\mathbf{T}$ improves the match of image A to image B, but consists only of such transformation components that do not change the appearance of image A. Consequently, the improvement of match cannot be detected by measuring the similarity $S(\mathbf{T}A, B)$. However, the same match as obtained by transforming image A using transformation $\mathbf{T}$ can also be obtained by transforming image B using the inverse transformation $\mathbf{T}^{-1}$. This corresponds to the other registration direction, where image B is registered to image A. The inverse transformation $\mathbf{T}^{-1}$ transforms image B into image $\mathbf{T}^{-1}B$, such that its appearance changes and the improvement of match can be detected by measuring similarity $S(A, \mathbf{T}^{-1}B)$. This illustration shows that measuring the similarity in different registration directions may detect different transformation components.

Asymmetry of similarity measurement reflects in asymmetry in external energy, such that external energy also depends on the registration direction. Consequently, when registering images in different registration directions, different registration results are obtained, i.e. different point to point image correspondence. These differences are also known as inverse inconsistency and indicate that at least one of the registration results is subjected to corresponding registration error. Similarly, differences between the registration results can also appear due to asymmetry of internal energy, defined by spatial deformation models [12]. However, in contrast to asymmetry of external energy, the asymmetry of internal energy can be physically and anatomically justified. Internal energy of some real tissue at certain transformation normally differ from internal energy of the same tissue at applied inverse transformation. Thus, when one image needs to be registered to the other one, incorrect results may be obtained by registering in the opposite registration direction, because of assuming different (incorrect) undeformed configuration.

One solution to the problem of inconsistency was proposed by Christensen and Johnson [13]. Their consistent image registration tends to improve the registration by jointly registering images in both registration directions and linking both processes with an additional consistency constraint. Another solution to this problem was presented by Cachier and Rey [12], who proposed inversion-invariant energies. Both solutions comprise measuring of similarity in both reg-

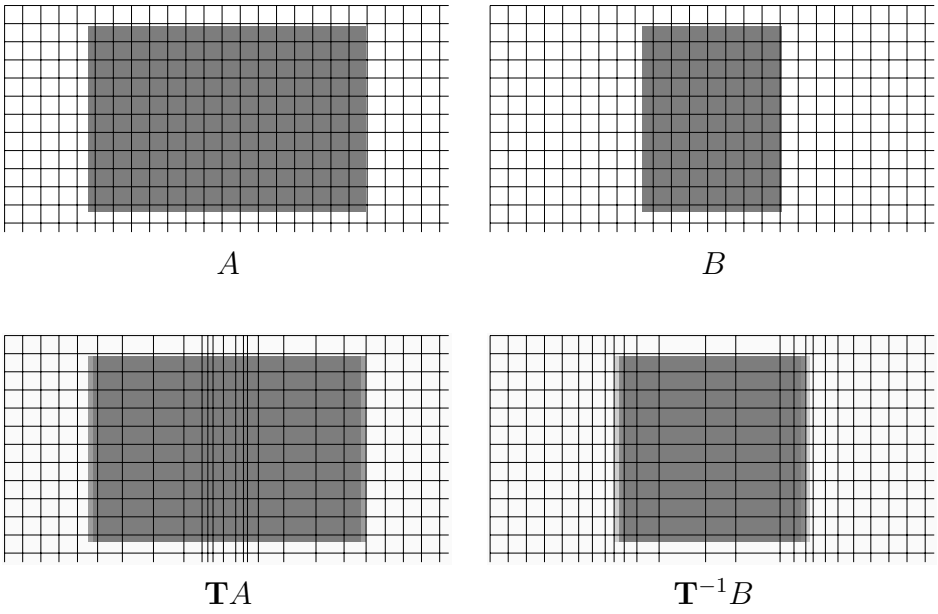


Figure 3.1: *Illustration of similarity measure asymmetry. Measurement of similarity in different registration directions may assess different transformation components. Similarity* $S(\mathbf{T}A, B)$ *between image* $\mathbf{T}A$ *and image* $\mathbf{T}A$ *do not differ from similarity* $S(A, B)$ *between the original (untransformed) images, which indicates that none of the transformation components present in transformation* $\mathbf{T}$ *can be assessed by measuring similarity in this registration direction. On the other hand, the same image correspondence as obtained by transforming image* A *with* $\mathbf{T}$ *can also be obtained in the other registration direction, by transforming image* B *with* $\mathbf{T}^{-1}$*. This transformation changes the appearance of image* B *and consequently, the similarity* $S(A, \mathbf{T}^{-1}B)$ *differs from similarity* $S(A, B)$*, which indicates that at least some transformation components can be assessed in this registration direction. This holds for any intensity based similarity measure. When using an overlap measure, the similarities are the following:* $S(\mathbf{T}A, B) = S(A, B) = 0.75$ *and* $S(A, \mathbf{T}^{-1}B) = 0.85$*.*

istration directions, which solves the problem of asymmetry of external energy. In addition, both approaches symmetrize the internal energies, which may not be always justifiable. Furthermore, both approaches require computation of inverse transformations, which is a difficult and computationally complex task.

We propose an alternative registration approach, which does not force the consistency, but improves the registration by establishing a symmetric image interdependence. The symmetric relationship enables the registration to gather more information about the image transformation by measuring the similarity in both registration directions, in order to symmetrize the external energy and thus improve the registration results. In addition, the proposed symmetric approach

does not restrict the properties of spatial deformation models, such that physically justifiable asymmetric deformation properties can be obtained as well. Finally, the proposed approach does not require computation of inverse transformation and is computationally effective.

3.2 Symmetric image registration

The symmetric registration approach treats both images involved in the registration process in the same manner. Both of the images share the same world coordinate system and can overlap. Both of the images may be modeled by spatial deformation models, possibly different ones, such that both of them may move and/or deform. Finally, the most distinctive feature of symmetric registration approach is interaction between the images. Images interact through forces, in accordance with the Newton's third law of motion. Forces on one image reflect in opposing forces on the other image, which forms the basis for the symmetry. The result of the interaction is transformation of the images that puts the whole system into the equilibrium state of minimal energy.

Let A and B be the images involved in the registration. Each of them is defined in its own coordinate system, $\mathbf{x}_A$ for image A and $\mathbf{x}_B$ for image B. Images are mapped to the world coordinate system $\mathbf{x}$ by transformations $\mathbf{T}_A$ and $\mathbf{T}_B$, such that $\mathbf{T}_A A$ and $\mathbf{T}_B B$ represent the transformed images as they appear in the world coordinate system $\mathbf{x}$. Here, the mappings of image coordinates to the world coordinates are:

$$\mathbf{x} = \mathbf{x}_A + \mathbf{T}_A(\mathbf{x}_A), \tag{3.1}$$

$$\mathbf{x} = \mathbf{x}_B + \mathbf{T}_B(\mathbf{x}_B), \tag{3.2}$$

where $\mathbf{T}_A(\mathbf{x}_A)$ and $\mathbf{T}_B(\mathbf{x}_B)$ denote displacements of points $A(\mathbf{x}_A)$ and $B(\mathbf{x}_B)$ from its initial (untransformed) position.

Following a widely used gradient descent optimization algorithm, external forces are defined as a gradient of image similarity $S(\mathbf{T}_A A, \mathbf{T}_B B)$. Forces F_A, which act on image A in order to match it with image B are then the following:

$$\mathbf{F}_A = \frac{\partial S(\mathbf{T}_A A, \mathbf{T}_B B)}{\partial \mathbf{T}_A}. \tag{3.3}$$

In addition to image A, image B also tends to improve the matching. Forces $\mathbf{F}_B$ are exerted on image B in order to improve matching of image B with image A:

$$\mathbf{F}_B = \frac{\partial S(\mathbf{T}_A A, \mathbf{T}_B B)}{\partial \mathbf{T}_B}. \tag{3.4}$$

The obtained forces, $\mathbf{F}_A$ and $\mathbf{F}_B$, which we call forward forces, are not symmetric and thus, they do not comply with physical laws, such that

$$\mathbf{F}_A(\mathbf{x}) \neq -\mathbf{F}_B(\mathbf{x}), \tag{3.5}$$

where $\mathbf{F}_A(\mathbf{x})$ and $\mathbf{F}_B(\mathbf{x})$ denote forces at the same world coordinates $\mathbf{x}$. Note that forces $\mathbf{F}_A$ and $\mathbf{F}_B$ act on different images. This asymmetry is also a source of inconsistency in the case of conventional non-rigid registration approach. Forces $\mathbf{F}_A$ and $\mathbf{F}_B$ are actually estimated by measuring similarity in different registration directions and so they are based on different information about image transformation. Each of the gradients of similarity as defined in Eq. (3.3) and Eq. (3.4) depend only on the gradient of one image. For example, if S is some global similarity measure, then forces $\mathbf{F}_A$ can differ from zero only at points with nonzero gradient of image A, while forces $\mathbf{F}_B$ differ form zero at points with nonzero gradient of image B. For illustration of the forward forces see Figure 3.2.

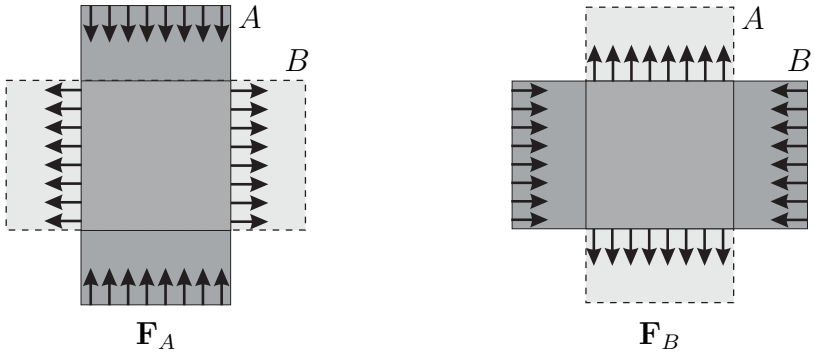


Figure 3.2: *Illustration of forces $\mathbf{F}_A$ acting on image A and $\mathbf{F}_B$ acting on image B when matching two rectangles. Note the asymmetry of the forces with respect to the registration direction, which is the main source of inconsistency of conventional registration procedures.*

However, the asymmetry does not affect the symmetric registration approach. In accordance with the Newton's third law of motion, each force exerted in one of the images reflects in another force of the same magnitude that act in the other image in the opposite direction. Thus, forces $\mathbf{F}_A$ reflect to forces $\mathbf{F'}_A$, which act on image B such that $\mathbf{F'}_A(\mathbf{x}) = -\mathbf{F}_A(\mathbf{x})$, and forces $\mathbf{F}_B$ reflect to forces $\mathbf{F'}_B$, which act on image A such that $\mathbf{F'}_B(\mathbf{x}) = -\mathbf{F}_B(\mathbf{x})$. The new forces are called reverse forces and are illustrated in Figure 3.3. The resultant forces that act on image A are the sum of forward and reverse forces:

$$\mathbf{F}^A(\mathbf{x}) = \mathbf{F}_A(\mathbf{x}) + \mathbf{F'}_B(\mathbf{x}) = \mathbf{F}_A(\mathbf{x}) - \mathbf{F}_B(\mathbf{x}), \tag{3.6}$$

and likewise the resultant forces on image B are

$$\mathbf{F}^B(\mathbf{x}) = \mathbf{F}_B(\mathbf{x}) + \mathbf{F'}_A(\mathbf{x}) = \mathbf{F}_B(\mathbf{x}) - \mathbf{F}_A(\mathbf{x}). \tag{3.7}$$

For illustration of resultant forces $\mathbf{F}^A$ and $\mathbf{F}^B$ see Figure 3.4.

Forces $\mathbf{F}^A$ and $\mathbf{F}^B$ are symmetric. They have the same magnitude but act on different images and in the opposite direction:

$$\mathbf{F}^A(\mathbf{x}) = -\mathbf{F}^B(\mathbf{x}). \tag{3.8}$$

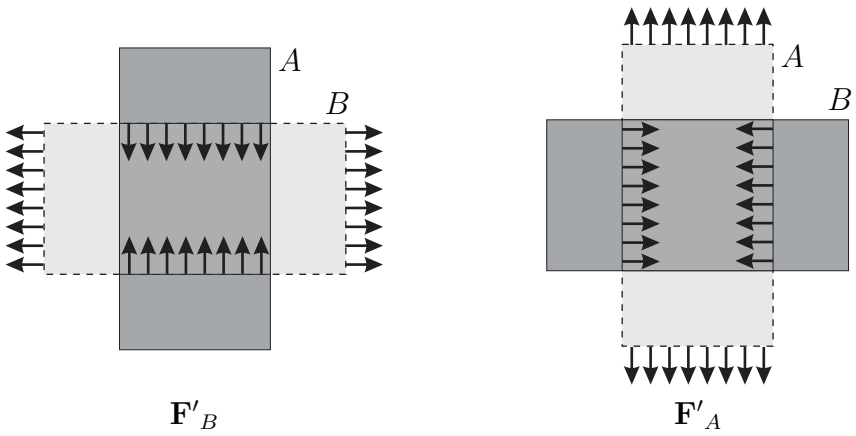


Figure 3.3: *Illustration of reverse forces* $\mathbf{F'}_B$ *acting on image A and* $\mathbf{F'}_A$ *acting on image B when matching two rectangles.*

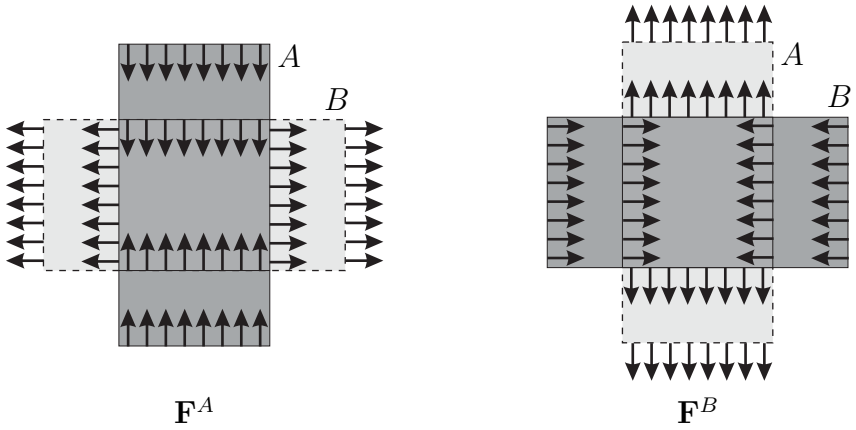


Figure 3.4: *Illustration of symmetric forces* $\mathbf{F}^A$ *acting on image A and* $\mathbf{F}^B$ *acting on image B when matching two rectangles. Note the increased amount of information available for transforming each of the images and the obtained symmetry* $\mathbf{F}^A = -\mathbf{F}^B$*, which is required for achieving the registration consistency.*

When the symmetric forces are used for the registration, each of the resulting transformations, $\mathbf{T}_A$ and $\mathbf{T}_B$ tend to correct all the image differences, i.e. those that can be detected in one and those that can be detected in the opposite registration direction. However, transformation $\mathbf{T}_A$ may be regularized with different spatial deformation model than transformation $\mathbf{T}_B$, which opens new possibilities of the registration. If the same spatial deformation model is used for both of the images, the registration is symmetric in all the aspects and registration results are absolutely consistent. However, it is often assumed that one of the images represents the undeformed configuration of the anatomy, and the other one is its deformed version. In this case the most correct results are obtained

when transforming only one of the images, while the other image should remain untransformed. Consequently, images must be modeled by two different spatial deformation models. The target image A, which should not get transformed, must be fixed (modeled as a rigid body and anchored to the coordinate system), while the source image B must be modeled using a suitable deformable model. Note that although the target image is fixed, the forces $\mathbf{F}_A$ still exist and they still contribute to the registration. Even if they cannot change the configuration of image A they still have influence on the transformation of source image B. Therefore, the advantages of the symmetric registration approach remain and improvement in registration correctness can be expected.

3.2.1 Implementation details

In our implementation of registration we assume that only one of the images needs to be transformed (image B), while the other image (A) is modeled as fixed. However, the forces that drive the registration are obtained using a symmetric approach, such that information of both registration directions is used.

For measuring the quality of image match we use point similarity measures described in Chapter 2., which are capable of estimating the similarity for individual image point pairs $(A(\mathbf{x}_1), B(\mathbf{x}_2))$. Instead of optimizing the similarity of the whole images, the improvement of image match is searched by optimizing similarities of individual image voxels. This is convenient for high dimensional registration, where transformation components represent displacements of individual voxels. Consequently, estimation of external forces, which are also estimated for individual voxels, can be simplified, as each force $\mathbf{F}(\mathbf{x})$ depends only on one transformation component $\mathbf{T}(\mathbf{x})$ and not on the whole transformation (of the whole image) $\mathbf{T}$.

Transformation $\mathbf{T}_B$, which transforms image B into image $\mathbf{T}_B B$, moves each point from its original (untransformed) position $\mathbf{x} = \mathbf{x}_B$ to a new position $\mathbf{x} = \mathbf{x}_B + \mathbf{T}_B(\mathbf{x}_B)$. Consequently, point at coordinate $\mathbf{x}$ in image $\mathbf{T}_B B$, denoted $\mathbf{T}_B B(\mathbf{x})$, is actually point $\mathbf{x}_B$ in image B (according to image coordinate system), denoted $B(\mathbf{x}_B)$:

$$\mathbf{T}_B B(\mathbf{x}) = B(\mathbf{x}_B), \tag{3.9}$$

Point $\mathbf{T}_B B(\mathbf{x})$ gets matched with point $\mathbf{T}_A A(\mathbf{x})$, but because image A is fixed ($\mathbf{T}_A(\mathbf{x}_A) = 0$) its image coordinates equal world coordinates $\mathbf{x}_A = \mathbf{x}$, see eq.(3.1). Thus, at coordinate $\mathbf{x}$, image point $\mathbf{T}_A A(\mathbf{x})$ matches with image point $\mathbf{T}_B B(\mathbf{x})$, such that

$$\mathbf{T}_A A(\mathbf{x}) = A(\mathbf{x}) = A(\mathbf{x}_B + \mathbf{T}_B(\mathbf{x}_B)). \tag{3.10}$$

At this position the forward force $\mathbf{F}_B(\mathbf{x})$ is the following:

$$\begin{aligned}\mathbf{F}_B(\mathbf{x}) &= \frac{\partial S(\mathbf{T}_A A(\mathbf{x}), \mathbf{T}_B B(\mathbf{x}))}{\partial \mathbf{T}_B} \\ &= \frac{\partial S(A(\mathbf{x}_B + \mathbf{T}_B(\mathbf{x}_B)), B(\mathbf{x}_B))}{\partial \mathbf{T}_B(\mathbf{x})}.\end{aligned} \tag{3.11}$$

The gradient of point similarity can be computed numerically, using similarities that correspond to point displacements $\Delta\mathbf{T}_B(\mathbf{x}) = [-\varepsilon, 0, +\varepsilon]^3$ and form a $3\times3\times3$ point similarity window, see Figure 3.5. Similarities for points in image A that are not positioned on the image grid are interpolated from the similarities of neighboring grid points. Further estimation of forces, i.e. gradients of similarity, follows Bajcsy and Kovačič [4].

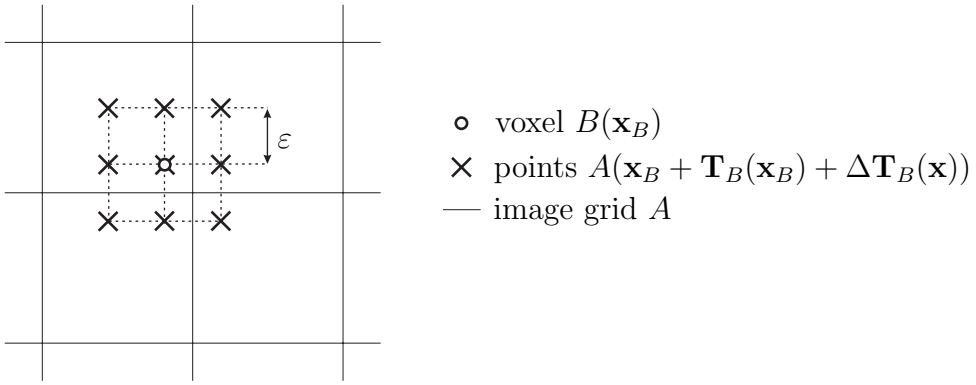


Figure 3.5: *2D illustration of points used for estimation of forward forces* $\mathbf{F}_B$.

For the computation of reverse forces $\mathbf{F'}_A = -\mathbf{F}_A$, a gradient of similarity with respect to transformation $\mathbf{T}_A$ must be obtained. Although the $\mathbf{T}_A$ is in reality zero (image A is fixed) we need to compute how the similarity would change when registering in the opposite direction. In this case point $A(\mathbf{x}_B + \mathbf{T}_B(\mathbf{x}_B))$ would no more match with point $B(\mathbf{x}_B)$, but with some other point in image B, displaced from coordinate $\mathbf{x}_B$ for some displacement $\mathbf{u}$. To avoid interpolation of image B, which is difficult because image B may be deformed, the gradient of similarity is numerically estimated at applied small displacements $\mathbf{u}$ instead of computing it at applied displacements $\mathbf{T}_A(\mathbf{x})$. Thus, the derivation is performed in the coordinate system of image B instead of using the world coordinate system $\mathbf{x}$, see Figure 3.6. The reverse force $\mathbf{F}_A(\mathbf{x}_B)$ estimated in this point according to coordinate system of image B is the following

$$\mathbf{F}_A(\mathbf{x}_B) = \left.\frac{\partial S(A(\mathbf{x}_B + \mathbf{T}_B(\mathbf{x}_B)), B(\mathbf{x}_B + \mathbf{u}))}{\partial \mathbf{u}}\right|_{\mathbf{u}=0}. \tag{3.12}$$

However, the final reverse forces must be defined according to the world coordinate system $\mathbf{x}$. The local transformation between the coordinate systems follows

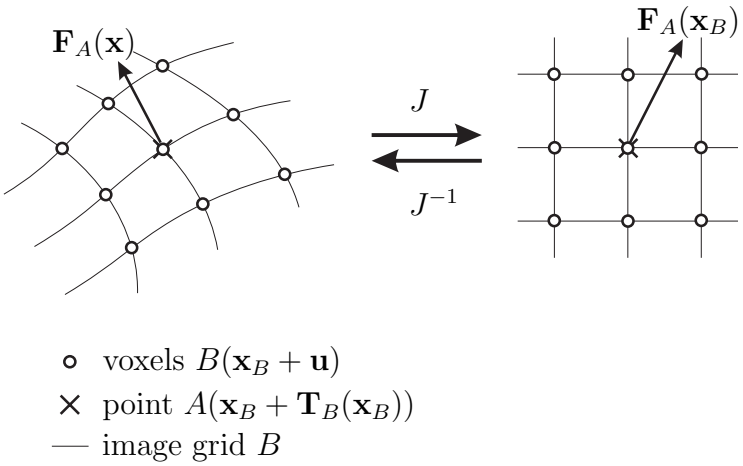


Figure 3.6: *2D illustration of points used for numerical computation of reverse forces. The numerical derivation is performed in the coordinate system of image B and the result is mapped to the world coordinate system using a Jacobian matrix J.*

Eq. (3.2) and equals the Jacobian matrix J:

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}_B} = J = I + \frac{\partial \mathbf{T}(\mathbf{x}_B)}{\partial \mathbf{x}_B}, \tag{3.13}$$

$$\frac{\partial \mathbf{x}_B}{\partial \mathbf{x}} = J^{-1}. \tag{3.14}$$

Thus, the reverse forces defined in a world coordinate system are

$$\mathbf{F}_A(\mathbf{x}) = J^{-1}\mathbf{F}_A(\mathbf{x}_B). \tag{3.15}$$

When both sets of forces, $\mathbf{F}_B$ and $\mathbf{F}_A$, are obtained, the resultant forces $\mathbf{F}^B$, used for registering image B to image A, are obtained as (3.7):

$$\mathbf{F}^B(\mathbf{x}) = \mathbf{F}_B(\mathbf{x}) - \mathbf{F}_A(\mathbf{x}). \tag{3.16}$$

3.3 Results

Three sets of experiments were performed to demonstrate the symmetric registration approach and compare it with mono-directional approaches. The experiments involve registration of simple objects, recovering of synthetic deformations, and interpatient registration of real images of human head.

Each pair of images (A, B) used in the experiments was registered using the three different force estimation approaches, one symmetric and two unidirectional

approaches, and in both registration directions. In total, six transformations were obtained for each image pair. Registration of image B to image A results in transformations $\mathbf{T}_F$ for standard registration with forward forces $\mathbf{F}_B$, $\mathbf{T}_R$ for registration based on reverse forces $\mathbf{F'}_A$, and $\mathbf{T}_S$ for symmetric registration approach driven by forces $\mathbf{F}^B$. When registration is performed in the opposite direction, registering image A to image B, transformations $\mathbf{T'}_F$, $\mathbf{T'}_R$ and $\mathbf{T'}_S$ are obtained, see Figure 3.7.

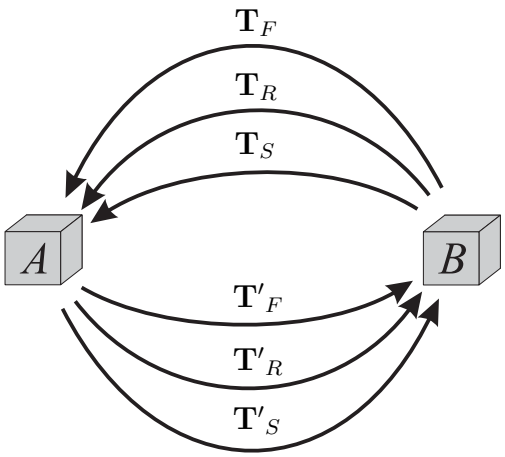


Figure 3.7: *Transformations obtained by registering images A and B, with respect to the method used for estimating forces (forward, reverse and symmetric) and the registration direction.*

All the experiments were performed using the system for high-dimensional multi-modality registration, which is described in Chapter 6. and based on multi-modality point similarity measure S_{UH}, see Chapter 2. Eq. 2.57. The registration was regularized using a combined elastic-incremental model, see Chapter 4., which follows the implementation of linear elasticity proposed by Bro-Nielsen [5].

Three different methods were used to assess the registration quality. First, when the correct transformation T_0 was known, the registration quality was measured directly by computing RMS displacement error E_{RMS},

$$E_{RMS}(\mathbf{T}) = \sqrt{\frac{1}{N}\sum_{\mathbf{x}} \left(\mathbf{T}(\mathbf{x}) - \mathbf{T}_0(\mathbf{x})\right)^2}. \tag{3.17}$$

Here N is the number of image voxels. The second method measures the registration consistency by computing a RMS deviation of point correspondences defined by transformations $\mathbf{T}$ and $\mathbf{T'}$, which are obtained by registering the same images using the same method but in different registration directions.

$$C_{RMS}(\mathbf{T}, \mathbf{T'}) = \sqrt{\frac{1}{N}\sum_{\mathbf{x}} \left(\mathbf{T}(\mathbf{x}) - \mathbf{T'}^{-1}(\mathbf{x})\right)^2} \tag{3.18}$$

$\mathbf{T}'^{-1}$ denotes the inverse of transformation $\mathbf{T}'$, such that it forms the same image correspondence in the other registration direction. The last method used for assessing the registration quality was measuring of image similarity. This method is less appropriate [89, 68] as it can only judge transformation components that change the image appearance. For our experiments correlation coefficient $CC(A, \mathbf{T}B)$ was used, see Eq. (2.3) and [39].

3.3.1 Registration of simple objects

Two images of simple objects were generated, a rectangular prism for image A and a sphere for image B. Central image slices are shown in Figure 3.8.

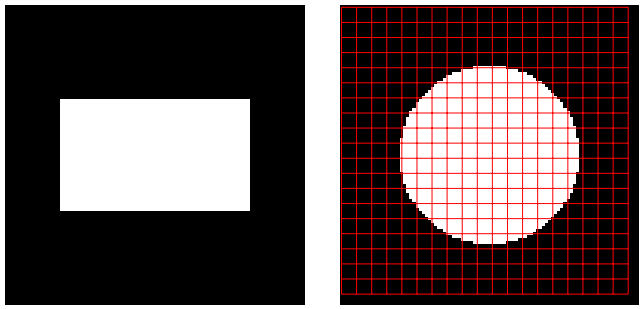


Figure 3.8: *The central slices of the simple 3D images, representing a rectangular prism (image A) and a sphere (image B).*

After registering the images with all three registration approaches in both directions, the results were compared by measuring consistency C_{RMS} and similarity of registered images CC. As the ideal transformation $\mathbf{T}_0$ is not known in this case, the registration error E_{RMS} can not be obtained. The results are shown in Figure 3.9. Although all three resulting images look very similar, the consistency measure C_{RMS} indicates the advantage of the symmetric approach. The results of the other two methods are worse, especially when using only the reverse forces (transformation T_R). However, different conclusions could be drawn when observing the image similarity. Here, the method based on the reverse forces gives the best result, better than the symmetric approach. To find out which of the approaches is more correct we performed the second experiment, based on recovering synthetic deformations.

3.3.2 Recovering synthetic deformations

In this experiment Brainweb [42] simulated images of human head were used. Firstly, the original MRI-T1 image was used for A and its synthetically deformed

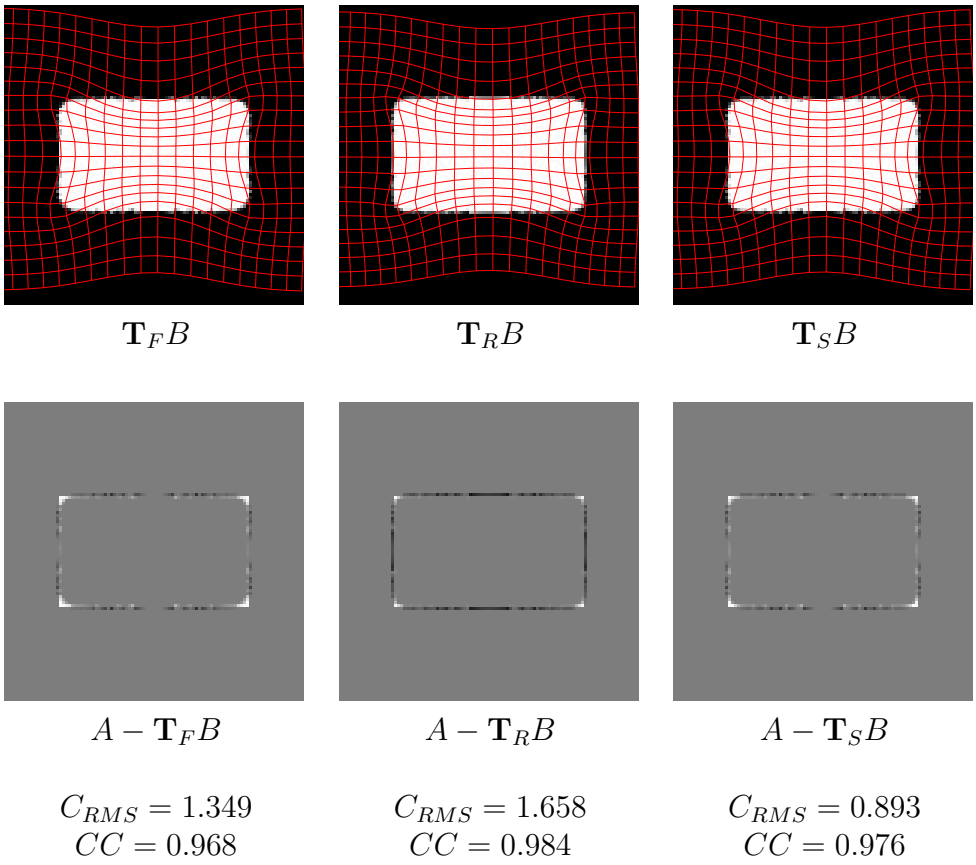
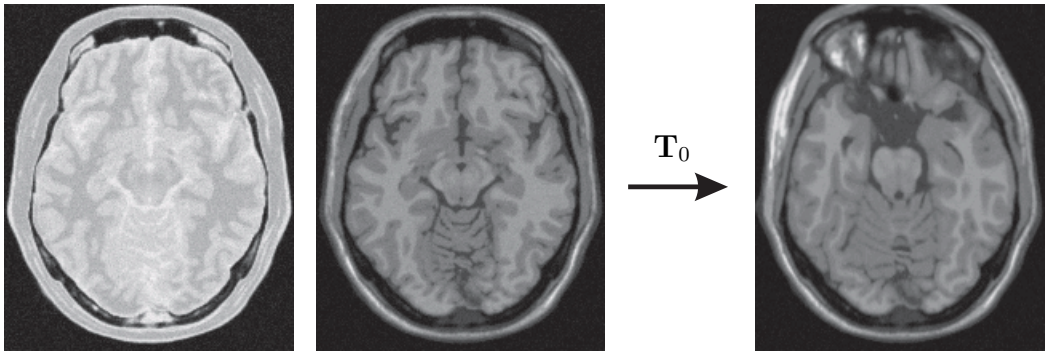


Figure 3.9: *Results of registering the simple image pair using different methods for estimating external forces. The first row represents the resulting images* $\mathbf{T}B$*, and the second row shows the difference to the target image* A*. Below the images, consistency results* C_{RMS} *and final image similarities* S_{MAD} *are shown for each of the methods.*

version for B. Secondly, we performed a multi-modality registration by using MRI-PD image as a target A, while B remained the same synthetically deformed MRI-T1 image, see Figure 3.10.

Five different synthetic deformations $\mathbf{T}_0$ were used, each of them generated as a sum of five three-dimensional Gaussian functions with randomly selected parameters, where the amplitude was in the range of 0 to 26mm, and standard deviation was in the range of 5 to 50mm.

The synthetically generated transformation $\mathbf{T}_0$ enabled the evaluation of registration correctness E_{RMS}, the consistency C_{RMS} and image similarity CC. Because the original images MRI-T1 and MRI-PD were registered, measure CC and original MRI-T1 image were used also for evaluation of the multi-modality registration results ($\mathbf{T}B$). The results are tabulated in Table 3.1. In all the



MRI-PD
A (multi-modality)

MRI-T1
A (mono-modality)

MRI-T1
B

Figure 3.10: *Images used for the experiment based on recovering synthetic deformations. Image A was an original untransformed MRI-T1 image (mono-modality registration) or MRI-PD image (multi-modality registration). Image B was in all the cases generated by deforming the original MRI-T1 image.*

cases the symmetric approach performed the best regarding to the registration correctness and registration consistency, while measuring of final image similarity gives similar results for all three registration approaches (considering the average initial image similarity $S_0 = 0.841$). It is also evident that mono-modality and multi-modality registration produce practically identical registration results.

3.3.3 Registration of real interpatient data

In the last experiment we tested how the symmetric registration approach performs in case of real medical images and complex interpatient transformations. Specifically, six real MRI-T1 images of human head were used, such that one of them always served as image A and the other five images were used as image B. Thus, altogether there were five image pairs and each one of them was registered six times, by all three different methods and in both directions. Example images are shown in Figure 3.11.

The registration results were evaluated by measuring the consistency C_{RMS} and image similarity CC, while registration error E_{RMS} cannot be measured due to the unknown ideal transformation $\mathbf{T}_0$. The results are tabulated in Table 3.2. One can observe that consistency of symmetric registration is in all cases much better than the consistency of the other two methods. Consequently, it can be expected that the results of the symmetric registration are also more correct. However, it is also evident that the symmetric approach does not improve the final image similarity, which is similar for all the three methods. This indicates

Table 3.1: *Results of recovering the synthetic deformations.*

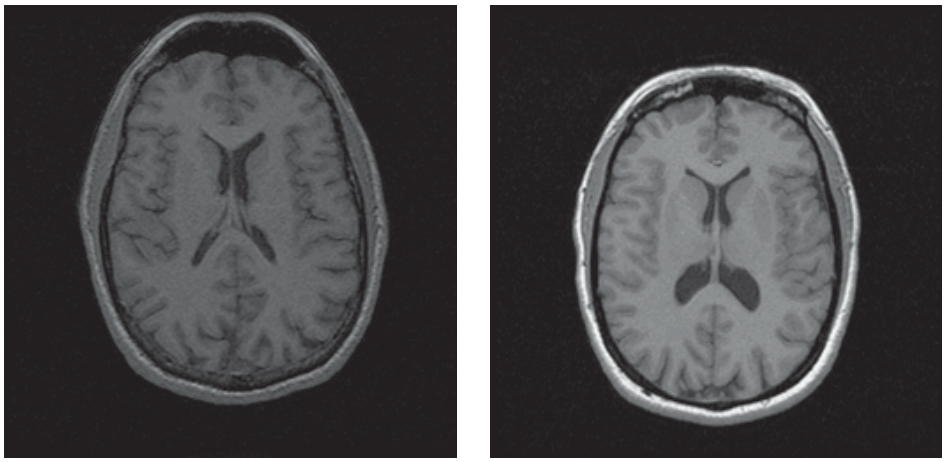
Experiment	Method	Mono-modality			Multi-modality		
		E_{RMS}	C_{RMS}	CC	E_{RMS}	C_{RMS}	CC
$\mathbf{T}_0^{(1)}$	$\mathbf{T}_F$	1.267	0.548	0.974	1.302	0.854	0.973
	$\mathbf{T}_R$	1.382	1.570	0.976	1.876	2.050	0.974
	$\mathbf{T}_S$	1.052	0.322	0.977	1.108	0.400	0.976
$\mathbf{T}_0^{(2)}$	$\mathbf{T}_F$	1.520	0.908	0.911	1.241	1.020	0.993
	$\mathbf{T}_R$	1.358	1.885	0.995	1.791	2.086	0.995
	$\mathbf{T}_S$	1.242	0.356	0.994	1.093	0.326	0.995
$\mathbf{T}_0^{(3)}$	$\mathbf{T}_F$	1.413	0.785	0.988	1.716	0.819	0.986
	$\mathbf{T}_R$	1.225	2.066	0.994	1.760	2.200	0.992
	$\mathbf{T}_S$	0.981	0.768	0.991	1.040	0.672	0.992
$\mathbf{T}_0^{(4)}$	$\mathbf{T}_F$	1.043	0.305	0.995	1.166	0.565	0.994
	$\mathbf{T}_R$	1.233	1.509	0.995	1.801	2.001	0.994
	$\mathbf{T}_S$	0.919	0.192	0.995	1.037	0.255	0.994
$\mathbf{T}_0^{(5)}$	$\mathbf{T}_F$	1.580	1.005	0.990	1.661	1.103	0.989
	$\mathbf{T}_R$	1.501	1.663	0.993	1.970	2.203	0.992
	$\mathbf{T}_S$	1.425	0.491	0.992	1.521	0.509	0.992
average	$\mathbf{T}_F$	1.578	0.916	0.987	1.417	0.872	0.987
	$\mathbf{T}_R$	1.489	1.731	0.991	1.840	2.108	0.989
	$\mathbf{T}_S$	1.330	0.510	0.990	1.160	0.433	0.990

that similarity measures are not capable of detecting differences between transformations obtained by different methods.

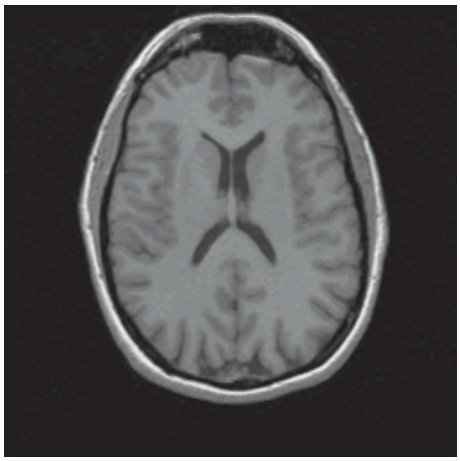
3.3.4 Discussion

All the results show that symmetric registration approach performs better in terms of consistency than standard registration approaches, thereby indicating more correct registration. Furthermore, the results of recovering synthetic deformations prove that the symmetric registration also improves the registration correctness. However, the improvement of registration error is smaller than the improvement of consistency, which indicates that some transformation components cannot be detected by measuring image similarity in any of the registration directions. Note that gradients of similarity and external forces still appear only for image points with nonzero gradient of intensity, which renders the exact registration of homogeneous image regions impossible. To improve the registration correctness of these regions an appropriate spatial deformation model must be used.

The results also show that the final image appearance in general becomes most



A B



$\mathbf{T}B$

Figure 3.11: *An example of interpatient registration of real MRI-T1 images of human head. Image A is a target used for registering image B, and $\mathbf{T}B$ is the registered image.*

similar to the target image when only the reverse forces are used. One possible explanation is that the reverse forces actually optimize the image similarity in the same registration direction as observed after the images are registered. As such, all the external energy is used exclusively for recovering the transformation components that improve appearance of the final deformed image, and not for other transformation components, which would also improve the registration correctness.

While in our implementation the reverse forces appear at the intensity gradients of the source image, the forward forces appear at gradients of the target

Table 3.2: *Results of interpatient registration of real images of head.*

Experiment	Method	C_{RMS}	CC
MRI01	$\mathbf{T}_F$	3.090	0.788
	$\mathbf{T}_R$	4.859	0.839
	$\mathbf{T}_S$	1.384	0.818
MRI02	$\mathbf{T}_F$	3.145	0.851
	$\mathbf{T}_R$	5.136	0.878
	$\mathbf{T}_S$	0.963	0.875
MRI03	$\mathbf{T}_F$	3.155	0.859
	$\mathbf{T}_R$	4.142	0.864
	$\mathbf{T}_S$	1.062	0.870
MRI04	$\mathbf{T}_F$	3.410	0.874
	$\mathbf{T}_R$	5.323	0.880
	$\mathbf{T}_S$	1.023	0.883
MRI05	$\mathbf{T}_F$	2.951	0.863
	$\mathbf{T}_R$	4.747	0.881
	$\mathbf{T}_S$	0.911	0.880
average	$\mathbf{T}_F$	3.150	0.847
	$\mathbf{T}_R$	4.841	0.868
	$\mathbf{T}_S$	1.068	0.865

image.During the registration the image match changes and the later forces appear at different coordinates with respect to the image that is being registered ($\mathbf{T}B$). This contributes to the registration consistency of forward force estimation method (comparing to the results for the reverse forces).

The consistency of the symmetric registration approach is considerably better than consistency of the other two methods, although not as good as one would expect. We have to emphasize that differences between results obtained in different registration directions are also caused by spatial deformation model, or to be more specific, by using different initial image configuration for the undeformed state of the anatomy [12]. Because properties of real tissues are not symmetric, two transformations that are inverses of each other require different deformation energies. Therefore, more correct registration results can be obtained when the correct undeformed configuration of anatomy is presumed. This is also a source of differences that contribute to the inverse inconsistency. However, note that such inconsistency may also be desired, as it allows more correct registration if the registration direction is selected properly.

In contrast to the approach proposed by Christensen and Johnson [13] the symmetric registration does not force the consistency. Instead, it improves the registration in one direction, without performing it in the opposite direction. As such it allows realistic asymmetric assumptions about the undeformed configu-

ration of anatomy and/or about the spatial deformation models. Furthermore, in comparison to the approach proposed by Cachier and Rey [12] it only symmetrizes the *similarity energy* while it does not constrain the *regularization energies*. When the symmetrization of the *regularization energies* is also desired it can be obtained by using the same spatial deformation model for both of the images. Finally, the implementation of the symmetric registration is computationally efficient and does not require computation of inverse transformations.

3.4 Conclusion

Symmetric image registration is a new aspect of the registration process, which is physically oriented and uses general physical laws. This solves the problem of asymmetry of similarity measurement, which is the main source of inconsistency and one of the sources of the registration errors.

Three different experiments were performed to demonstrate the symmetric registration approach and compare it with two standard unidirectional approaches. The results prove that the symmetric registration approach does improve the registration consistency as well as the registration correctness. In addition, the computational cost of the symmetric approach is still relatively low, as it does not require computation of inverse transformations.

The symmetric approach in general allows both images to be modeled by suitable deformable models and thus both of them could actually deform. Such approach could better suit to certain registration tasks when both of the images actually represent the deformed anatomy. In addition to this generalization, it would also be possible to use more than two images, which would interact with forces at the same time.

Finally, note that the registration process is completely symmetric only if both of the images are modeled using the same spatial deformation model. When this is not the case, the registration error depends on the selection of registration direction and some inconsistency appears due to the different assumptions in different registration directions.

4. Spatial Deformation Models

Spatial deformation models regularize registration by introducing spatial independence of neighboring image points, and thus preventing anatomically unlikely transformations. Several approaches regularize registration by modeling physical material properties, e.g. elasticity or viscosity. Such physical modeling is in general expected to aid to more correct transformation and thus better registration results. However, the deformation depends not only on spatial deformation model, but also on external forces, which drive the registration. Thus, in order to obtain realistic results, the external forces should equal the forces that have actually deformed the anatomy. Unfortunately, this is never the case because images do not contain sufficient information about the deformation of anatomy, and the external forces act only on information rich image regions (edges). Spatial deformation models have to be designed carefully to obtain realistic transformations from such unrealistic external forces. In this chapter we focus on spatial deformation models based on convolution. We analyze advantages and disadvantages of most commonly used spatial deformation models, i.e. elastic model, viscous model and incremental model, and propose a new combined model, which combines advantages of the elastic and the incremental model. Furthermore, we compare different models and two common convolution kernels: an elastic kernel and a Gaussian kernel.

4.1 Convolution models

Deformation of materials is in physics described by various physical laws. The same deformation properties as characterize real materials are usually expected from spatial deformation models that are used for non-rigid registration. However, in order to exactly model the behavior of realistic materials multiple physical laws must be employed at the same time. In practice, such an exact modeling is not required and spatial deformation models follow only a single, the most characteristic physical law, i.e. elasticity [4, 13, 33, 34] or viscosity [25, 14, 6]. Furthermore, each characteristic behavior of deformations can be linearized and thus the deformation properties of real materials can be approximated by linear models, which are computationally more efficient. Such idealized spatial de-

formation models can be implemented using a couple of approaches. Initially, they were modeled exclusively by implementing corresponding partial differential equations [4]. Later a finite element approach [26] and a convolution approach [6] were proposed. In this work we follow the convolution approach, proposed by Bro-Nielsen [5, 6]. The idea of this approach is that every linear model can be implemented by convolution filtering, where the filter kernel equals the impulse response of the deformable media. Such regularization can be applied to voxel displacements or to voxel velocities. Thus, common spatial deformation models differ only in two aspects: according to the filter kernel and according to the data that is being filtered.

4.1.1 Elastic models

The characteristic behavior of the elastic materials is that they deform due to the external forces and when the forces are retracted the elastic materials return back into the undeformed configuration. Following the work of Bro-Nielsen [5], this relation between forces $\mathbf{F}$ and transformation $\mathbf{T}$ can be described using a convolution approach:

$$\mathbf{T} = \mathbf{G} * \mathbf{F}, \tag{4.1}$$

where $\mathbf{G}$ stands for the spatial convolution filter with kernel that equals the impulse response of the deformable media. However, this equation cannot be directly used for image registration purposes, because the forces that would bring the images into register cannot be directly computed from the mismatched images. External forces, which drive the registration, are only an estimate how the forces should act to *improve* the image match.

In order to solve this problem we separate the spatial deformation model into two parts. The first part follows the Hooke's law to compute unregularized displacement of image points, while the second part regularizes them according to the impulse response of the deformable media to obtain the final transformation $\mathbf{T}$, see Fig. 4.1. In accordance with this we separate the convolution filter $\mathbf{G}$ into two terms, a filter gain k_F and a normalized kernel $\mathbf{G}_E$, such that

$$\mathbf{G} = k_F \mathbf{G}_E, \tag{4.2}$$

$$\int \mathbf{G}_E(\mathbf{x}) d\mathbf{x} = \sum_{\mathbf{x}} \mathbf{G}_E(\mathbf{x}) = 1. \tag{4.3}$$

Consequently, the elastic model (4.1) can be rewritten in the following way:

$$\mathbf{T} = \mathbf{G}_E * (k_F \mathbf{F}) = \mathbf{G}_E * \mathbf{T}_F, \tag{4.4}$$

where $\mathbf{T}_F$ is a transformation, i.e. displacements of independent image points due to the applied force, as follows from the Hooke's law, $\mathbf{T}_F = k_F \mathbf{F}$. Convolution

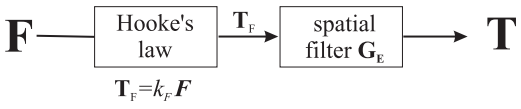


Figure 4.1: *Elastic model separated into the Hooke's law, which maps external forces into a transformation (displacements) of independent image points, and spatial regularization filter G_E, which models interdependencies between the points to regularize the transformation.*

filter $\mathbf{G}_E$ models interdependencies of image points, which imposes certain spatial deformation properties and thus regularize the transformation.

External forces obtained during the registration procedure tend to improve the transformation iteratively. Consequently, the transformation in the t-th iteration $\mathbf{T}^{(t)}$ is a sum of the transformation obtained in the previous iteration $\mathbf{T}^{(t-1)}$ and the increment $\mathbf{T}_F^{(t)}$ that follows the Hooke's law, and furthermore, all together must be regularized by the spatial regularization filter, see the block scheme in Figure 4.2:

$$\mathbf{T}^{(t)} = (\mathbf{T}^{(t-1)} + \mathbf{T}_F^{(t)}) * \mathbf{G}_E \tag{4.5}$$

This can be rewritten in the following form:

$$\mathbf{T}^{(t)} = \mathbf{T}_F^{(t)} * \mathbf{G}_E + \mathbf{T}_F^{(t-1)} * \mathbf{G}_E^2 + \mathbf{T}_F^{(t-2)} * \mathbf{G}_E^3 + ... + \mathbf{T}_F^{(1)} * \mathbf{G}_E^t, \tag{4.6}$$

where $\mathbf{G}_E^n$ stands for n-times convolution with filter $\mathbf{G}_E$. By increasing the num-

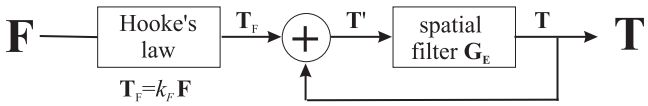


Figure 4.2: *A block scheme of the elastic model suitable for iterative non-rigid registration procedures.*

ber of convolution steps n the convolution kernel becomes wider and approaches towards averaging. Consequently, if the external forces do not exist ($\mathbf{F} = 0$), the model gradually returns back to the undeformed configuration. As forces in earlier iterations are regularized with wider kernels they contribute to more global matching, while forces in later iterations are regularized with narrower kernels and deal with more localized image mismatches. This is advantageous because the estimated forces do not act directly in the direction of the correct match, and thus include a local error. This is more obvious in initial registration iterations, when images are more mismatched, than in later registration iterations, when points used for estimating external forces are already close to their correct position. However, the elastic model also has one important disadvantage. The problem appears because external forces exist only if there is some

image mismatch, while they are necessary to maintain the deformed state of the image. Thus, local image discrepancies can never be registered absolutely correct as there is always some mismatch required to maintain the deformation. The mismatch is a systematic error and is larger for larger deformations. This makes the elastic model less suitable for large deformations. In such cases viscous fluid or incremental models are often used instead.

4.1.2 Viscous fluid models

Viscous fluid models were introduced in order to avoid the systematic error of elastic models and enable large deformations [14]. They model motion of highly viscous fluids. Instead of applying restriction to point displacements, they restrict point velocities $\mathbf{V}$:

$$\mathbf{V}^{(t)} = (\mathbf{V}^{(t-1)} + \mathbf{V}_F^{(t)}) * \mathbf{G}_V, \tag{4.7}$$

$$\mathbf{V}_F = k_F \mathbf{F}, \tag{4.8}$$

where $\mathbf{G}_V$ is a normalized kernel of convolution filter used to regularize point velocities.

$$\int \mathbf{G}_V(\mathbf{x}) d\mathbf{x} = \sum_{\mathbf{x}} \mathbf{G}_V(\mathbf{x}) = 1. \tag{4.9}$$

The transformation $\mathbf{T}$, i.e. the point displacements, is obtained from point velocities $\mathbf{V}$ by integration, which is in the discrete time (iteration) domain performed by summation:

$$\mathbf{T}^{(t)} = \sum_{n=1}^{t} \mathbf{V}^{(t)}. \tag{4.10}$$

The block scheme of the viscous fluid model based on convolution is shown in Fig. 4.3. The difference to the elastic model is only in the second summation, which serves as an integrator.

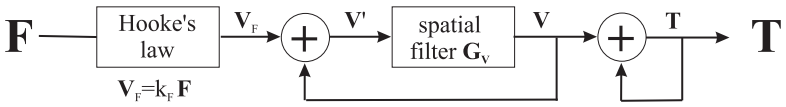


Figure 4.3: *A block scheme of the viscous fluid model.*

The viscous model considerably differs from the deformation properties of real tissues. When forces are retracted the material remains in the deformed configuration. Viscous model also enables very large deformations, even such that are not expected for real tissues. The disadvantage of this approach is in accumulation of registration error. External forces do not always act exactly in the direction of the correct match, which is most obvious in the initial registration steps. Components of external forces, which are perpendicular to the direction of

the correct match, represent an error, which accumulates in point velocities and point displacements, and therefore also reflects in the registration error. Thus, the ability to perform large image deformation is related to larger registration error. This is the main disadvantage of the viscous fluid model, which makes it less suitable for modeling deformations of real tissues.

4.1.3 Incremental models

Incremental models were introduced to avoid systematic error of the elastic approach. They are also based on the elasticity and assume that the total force $\mathbf{F}_0$, which is required for registering two images, can be obtained by summing the estimated external forces over all registration iterations:

$$\mathbf{F}_0 = \sum_t \mathbf{F}^{(t)}, \tag{4.11}$$

where t denotes the iteration number. Following the principle of linearity, the final transformation $\mathbf{T}$ can also be computed as a sum of partial transformations (displacements) $\mathbf{T}^{(t)}$

$$\mathbf{T} = k_F \mathbf{G}_E \sum_t \mathbf{F}^{(t)} = \sum_t \mathbf{G}_E \mathbf{T}_F^{(t)}, \tag{4.12}$$

and consequently,

$$\mathbf{T}^{(t)} = \mathbf{T}^{(t-1)} + \mathbf{G}_E \mathbf{T}_F^{(t)}, \tag{4.13}$$

A block scheme of incremental spatial deformation model is shown in Fig. 4.4.

Figure 4.4: *A block scheme of the incremental model.*

This kind of model has similar characteristics as viscous fluid models. When no external forces are applied it does not return into the undeformed state. Consequently, it accumulates the registration error, which is its main disadvantage. On the other hand it does allow large deformations and does not suffer from the systematic error known from the elastic models.

4.1.4 Combined elastic-incremental model

All the models described earlier have certain advantages and certain disadvantages. The elastic model provides physically and anatomically reasonable deformation properties and thus assures relatively low registration error for information

poor image regions, e.g. homogeneous image regions. However, it suffers from the systematic error, which is the most obvious in information rich regions, e.g. edges of anatomical structures may not perfectly overlap. On the other hand, deformation properties of the incremental model and the viscous model differ from deformation properties that are expected for most real tissues, which results in higher registration errors in information poor image regions. However, because these models do not suffer from the systematic error, they better register information rich image regions.

In order to combine the advantages of different spatial deformation models and improve the registration we devised a combined elastic-incremental model. The elastic part is expected to contribute to low registration error for information poor image regions, while the incremental part is expected to aid to matching of information rich image regions. The combined model would thus enable larger deformations and provide lower systematic error than the elastic model, and, on the other hand, improve the anatomical suitability of the incremental model. The obtained model consists of two convolution filters, where the first one, $\mathbf{G}_I$, follows the principle of the incremental model and regularizes transformation improvements $\mathbf{T}_F$, and the second filter, $\mathbf{G}_E$, represents the elastic properties and regularizes the overall transformation $\mathbf{T}$, see the block scheme in Figure 4.5:

$$\mathbf{T}^{(t)} = (\mathbf{T}^{(t-1)} + \mathbf{T}_F^{(t)} * \mathbf{G}_I) * \mathbf{G}_E \tag{4.14}$$

The first filter ($\mathbf{G}_I$) enables large deformations and precise registration while the second one ($\mathbf{G}_E$) serves to improve the linearity of the results.

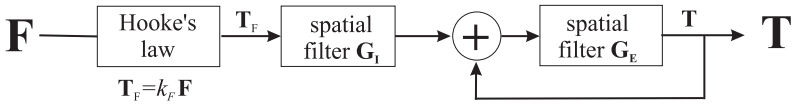


Figure 4.5: *A block scheme of the incremental model.*

The total regularization $\mathbf{G}$, i.e. the normalized impulse response of the combined model, suits to the convolution of both filter kernels:

$$\mathbf{G} = \mathbf{G}_I * \mathbf{G}_E. \tag{4.15}$$

If Gaussian filters are used, see Section 4.2, then the obtained standard deviation is

$$\sigma_{\mathbf{G}} = \sqrt{\sigma_{\mathbf{G}_I} + \sigma_{\mathbf{G}_E}}. \tag{4.16}$$

The behavior of the combined spatial deformation model depends on the ratio between regularization provided by each of the filters, such that in the extreme cases the incremental model is obtained when $\mathbf{G}_E = \delta$ and the elastic model is obtained when $\mathbf{G}_I = \delta$. Here δ is a Dirac's delta function. If both filters differ from δ, then the improvements are expected. A practical comparison between different settings of the filters is shown in Section 4.1.5.

4.1.5 Comparison

To test the combined elastic-incremental model and compare it with the elastic and the incremental model, we performed an evaluation study based on recovering synthetic deformations, see Section 5.3.1. The synthetic deformation was generated as a sum of six three-dimensional Gaussian functions with standard deviation ranging between 15 and 60 mm, such that the initial RMS displacement error was $e_{RMS} = 7.09$ mm and the maximal displacement error was $e_{max} = 16.84$ mm. The experiment was performed using two spatially aligned Brainweb [17, 42, 20] images of human head with voxel size $1 \times 1 \times 1$ mm. The first image, MRI-PD, was synthetically deformed and used as the target for registering the other image, MRI-T1. For more information about the system used for the registration see Chapter 6. In all the experiments a point similarity measure S_{UH} was used, see Section 2.4.3. Gaussian convolution kernels were used for the regularization, such that the standard deviation of the overall regularization was in all the cases the same, $\sigma_{\mathbf{G}} = 4.24$. The comparison was made between the results for nine different settings of the combined spatial deformation model, where one of the settings suited to the elastic model ($\sigma_{\mathbf{G}_I} = 0$,$\sigma_{\mathbf{G}_E} = 4.24$) and one to the incremental model ($\sigma_{\mathbf{G}_I} = 4.24$,$\sigma_{\mathbf{G}_E} = 0$). In all the cases the coefficient k_F was recomputed in each iteration, such that the increment $\mathbf{T}_F(\mathbf{x})$ was limited to the size of one image voxel. The comparison was made for three resolution levels of a multiresolution registration strategy. First, images subsampled by factor 4 were registered in resolution level 2, then the registration continued in level 1 with images subsampled by factor 2, and finally, the original images were registered in level 0. The best result of each level (with respect to maximal displacement error e_{max}) was used as the initial deformation for the next resolution level. The results are tabulated in Tables 4.1, 4.2 and 4.3.

Table 4.1: *Evaluation of the combined elastic-incremental spatial deformation model for the resolution level 2.*

$\sigma_{\mathbf{G}_I}$	$\sigma_{\mathbf{G}_E}$	e_{RMS}	e_{max}	CC
0.00	4.24	5.03	14.05	0.9206
0.50	4.21	4.72	13.60	0.9305
1.00	4.12	4.11	12.85	0.9444
2.00	3.74	3.16	11.42	0.9565
3.00	3.00	2.13	8.40	0.9657
3.74	2.00	1.37	4.68	0.9523
4.12	1.00	**1.30**	4.13	0.9532
4.21	0.50	1.17	4.64	0.9611
4.24	0.00	1.49	4.95	0.9523

In all three resolution levels the best results were obtained when both of the filters of the combined model were employed. Thus, the combined model

Table 4.2: *Evaluation of the combined elastic-incremental spatial deformation model for the resolution level 1.*

$\sigma_{\mathbf{G}_I}$	$\sigma_{\mathbf{G}_E}$	e_{RMS}	e_{max}	CC
0.00	4.24	3.19	11.32	0.9626
0.50	4.21	2.96	10.93	0.9671
1.00	4.12	2.44	9.98	0.9751
2.00	3.74	1.44	7.25	0.9819
3.00	3.00	0.80	3.74	0.9838
3.74	2.00	**0.55**	1.95	0.9837
4.12	1.00	0.59	2.41	0.9827
4.21	0.50	0.68	2.83	0.9806
4.24	0.00	0.73	3.09	0.9809

Table 4.3: *Evaluation of the combined elastic-incremental spatial deformation model for the resolution level 0.*

$\sigma_{\mathbf{G}_I}$	$\sigma_{\mathbf{G}_E}$	e_{RMS}	e_{max}	CC
0.00	4.24	1.51	6.19	0.9827
0.50	4.21	1.39	5.95	0.9840
1.00	4.12	1.16	5.33	0.9856
2.00	3.74	0.76	3.82	0.9869
3.00	3.00	0.50	2.62	0.9870
3.74	2.00	**0.39**	1.97	0.9863
4.12	1.00	0.43	1.78	0.9857
4.21	0.50	0.49	1.80	0.9854
4.24	0.00	0.51	1.92	0.9852

performed better than the incremental or the elastic model. The elastic model resulted in a large registration error due to small external forces, which were limited in order to limit the change of transformation in one registration iteration to the size of one image voxel. Consequently, the registration cannot result in larger deformations than those, that can be maintained by such small external forces. On the other hand, the incremental model also results in larger errors than the optimal combined model, which is due to non-linear relationship between external forces and required image transformation. The best results were obtained when both of the convolution filters were used, such that elastic and incremental properties were combined. In our case, the best results were obtained when the standard deviation of the incremental kernel was a bit larger than the standard deviation of the elastic kernel, but in general the optimal settings may depend on the type of deformations that need to be recovered.

4.2 Filter kernels

Kernels of the convolution filters define spatial deformation properties of the modeled material. Firstly, the kernel width defines stiffness of the model, such that wider kernels correspond to more stiff materials and narrower kernels correspond to more flexible materials. Secondly, the type of the kernel defines some other characteristics of the deformation, e.g. compressibility, isotropy, etc. Because the continuum mechanic non-rigid registration follows the principles of deformation of real-world materials, the elasticity and the viscosity are the most characteristic properties. When this two properties are modeled using a convolution approach, they require an elastic convolution kernel, which was proposed by Bro-Nielsen [6]. However, due to differences between real forces, which deform the anatomy, and the estimated forces that drive the registration, even such a realistic kernel does not guarantee good registration correctness. This gives rise to the use of simpler kernels, where the most frequently used one is a Gaussian kernel.

4.2.1 Elastic kernel

The elastic kernel was proposed by Bro-Nielsen and Gramkow in 1996 [6]. It enables modeling of linear elasticity and viscousity using the convolution approach. This approach speeds up the registration by the factor of 1000 with respect to the approach based on solving Navier-Stokes partial differentials equations [82]. For the details about the elastic kernel see [5, 6, 31]. An example of the elastic convolution kernel is shown in Figure 4.6.

4.2.2 Gaussian kernel

The Gaussian kernel can be regarded as a separable approximation to the elastic kernel. Due to the separability a three-dimensional Gaussian filter with the size N^3, which originally requires N^3 multiplications, reduces to three independent one-dimensional filters (one for each dimension), which requires only $3N$ multiplications. The substantial increase of speed is the main reason why Gaussian kernels have been extensively used. Here, let us mention only a popular 'demons-based' algorithm, proposed by Thirion in 1995 [80].

The main difference between the separable kernels, e.g. Gaussian, and more realistic elastic kernel is in the compressibility of the modeled material. Due to the independence of spatial dimensions the separable kernels do not provide control over the compressibility, and a longitudinal stretch does not induce a lateral shrink, see Figure 4.7.

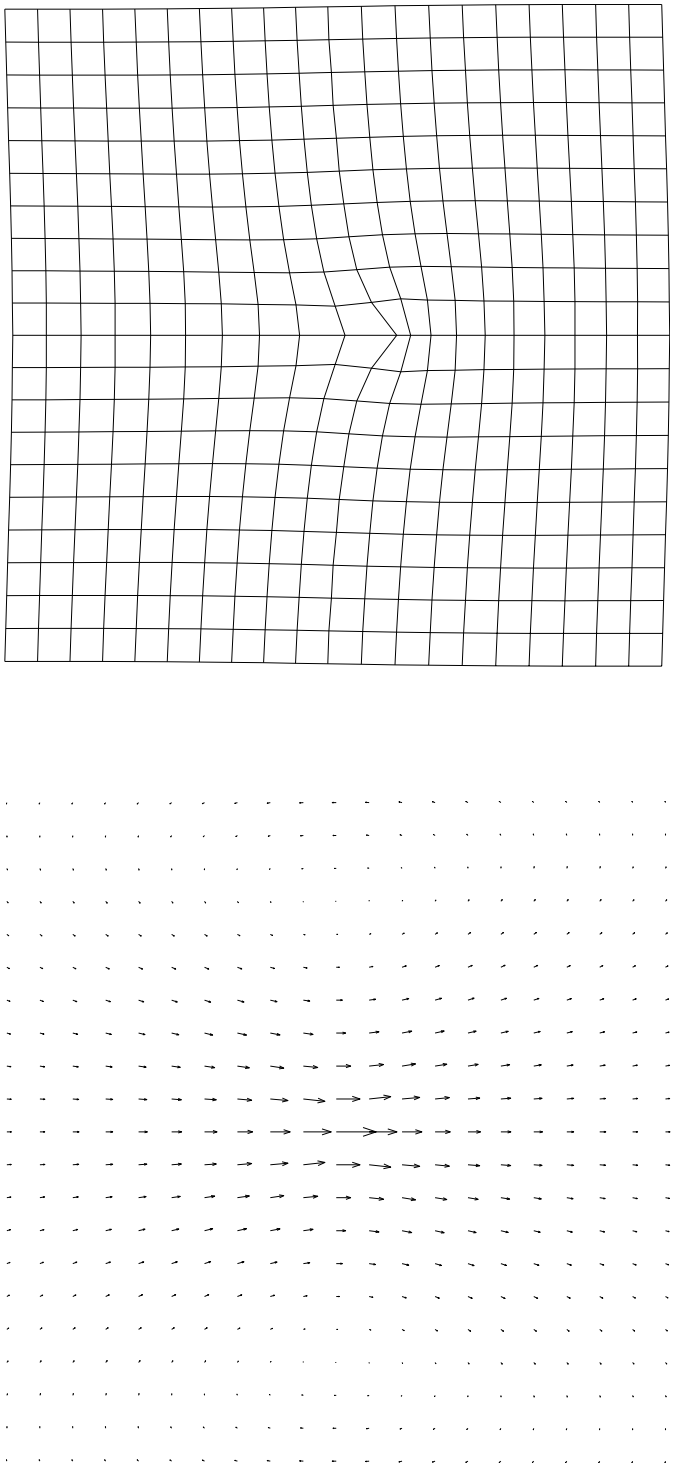


Figure 4.6: *An example of elastic kernel presented by a deformed grid (top) and by displacement vectors (bottom).*

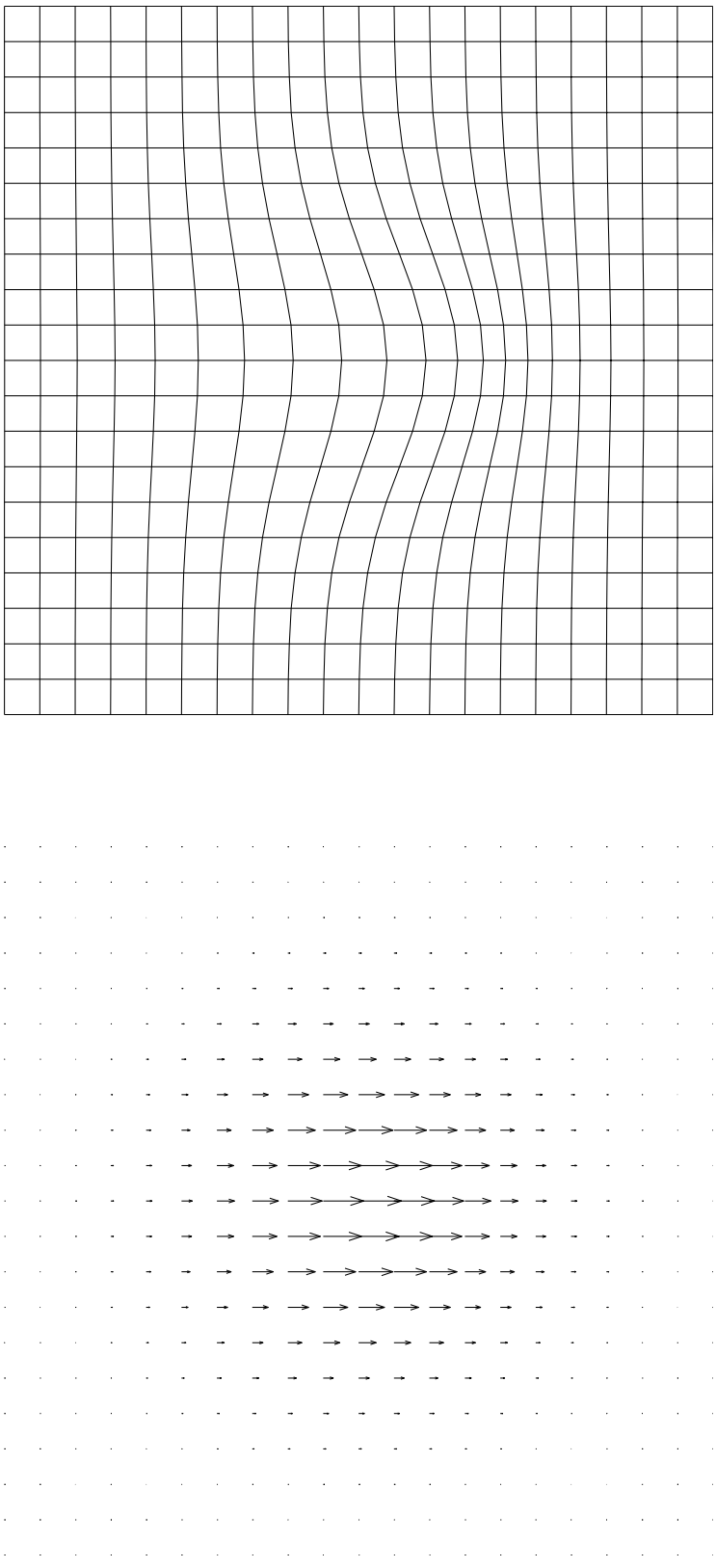


Figure 4.7: *An example of Gaussian kernel presented by a deformed grid (top) and by displacement vectors (bottom).*

4.2.3 Comparison

In order to compare the elastic kernel and the Gaussian kernel with respect to their suitability for image registration, we performed a simple 1D experiment, where the ability of the kernels to describe elastic deformations was observed. 1D experiments do not expose all the differences between the kernels, which are important in the case of 3D registration, because the relations between the spatial dimensions, which in the case of 3D registration define compressibility of the material, cannot be observed. Despite of this, the 1D experiment enables the analysis of registration error with respect to the position of forces, which in reality deform the registered anatomy. We have analyzed two possible situations: first, if the forces, which in reality deform the anatomy, act on edges of anatomical structures, and second, if the forces act on points inside the structures.

The situation, when the anatomy is deformed by forces that act on edges of anatomical structures, is illustrated in Figure 4.8. The figure represents an image of an edge between two tissue types, which are represented by two different intensity values. In our experiment both tissues have the same elastic deformation properties. If the tissues are subjected to a force F_0, they deform. The deformation (displacements of each point from the original position) is illustrated in Figure 4.9 with the thin solid line. Now, let us suppose that we register the original, undeformed image to the deformed one, using two different convolution kernels: first, the elastic kernel, which suits to the impulse response of the deformed material, and second, the Gaussian Kernel. The external force F_R estimated during the registration would ideally act on the same point as the force F_0, which was used to deform the tissues. Consequently, when the elastic kernel is used, the deformation can be perfectly recovered, see Figure 4.9, dashed line. However, if the Gaussian kernel is used, the deformation cannot be perfectly recovered, see Figure 4.9, dotted line. The point on the edge is correctly transformed, which means that the registration still manages to make the images look equal, but transformation deviates from the ideal one in points inside the homogeneous image regions.

In the second situation the force F_0 does not act on the edge between the tissues, but on some point inside a homogeneous region near the edge, as illustrated in Figure 4.10. The consequent deformation is similar to the deformation in the previous case, the only difference is in the position with respect to the edge, see Figure 4.11 thin solid line. However, when the original image is registered to the deformed one, the deformation in the homogeneous regions cannot be detected by measuring image similarity, such that the forces F_R act only on points with nonzero intensity gradient, which is mainly at the edges of anatomical structures, see Figure 4.10. Although the registration may deform images such that they look equal, the original deformation cannot be recovered correctly, not even when the realistic elastic kernel is used, see Figure 4.11, dashed line. However, if the Gaus-

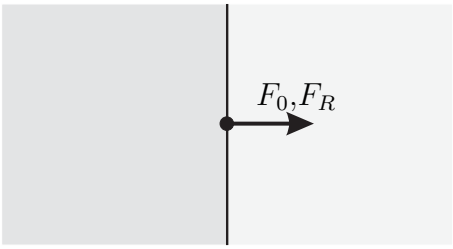


Figure 4.8: *An illustration of a force F_0 acting on the edge of an imaged structure. It can be expected that force F_R estimated during the registration acts on the same point as F_0, which enables good registration results.*

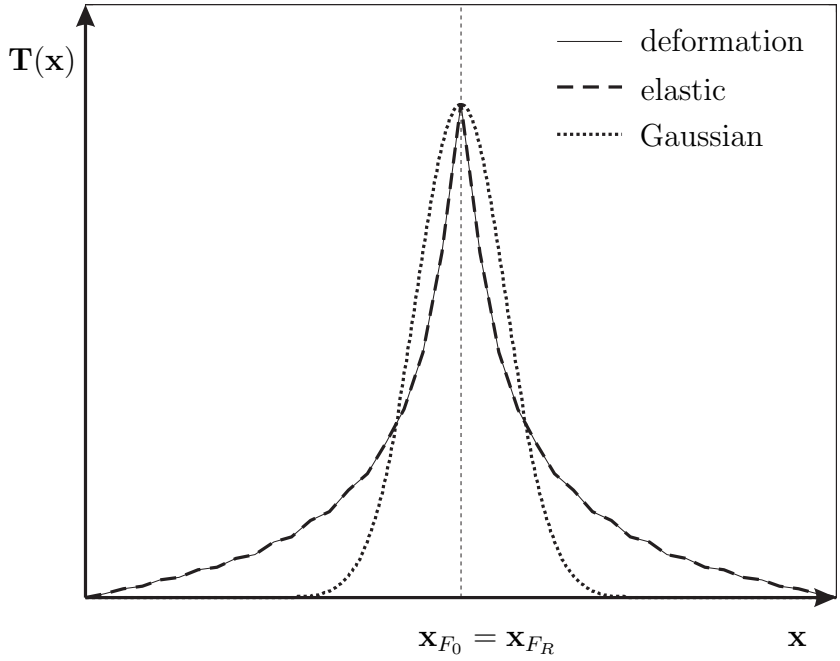


Figure 4.9: *Recovering of deformation of elastic material, deformed by the force F_0 that acts on the edge of an anatomical structure (the deformation is presented by the thin solid line). Registration may perfectly recover the deformation when the convolution kernel suits to the properties of the deformed media (dashed line). When different kernels are used, e.g. Gaussian, the registration cannot perfectly recover the deformation and some registration error exists (dotted line).*

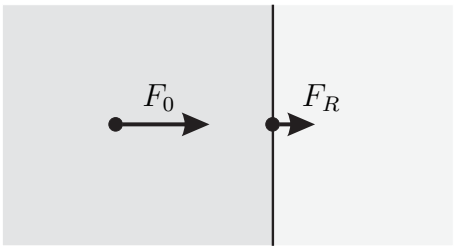


Figure 4.10: *An illustration of a force F_0 acting on a point in a homogeneous image region. The deformation, which it causes, cannot be recovered by registration, because the estimated forces F_R always act on the points with nonzero intensity gradient, which mostly correspond to edges of anatomical structures.*

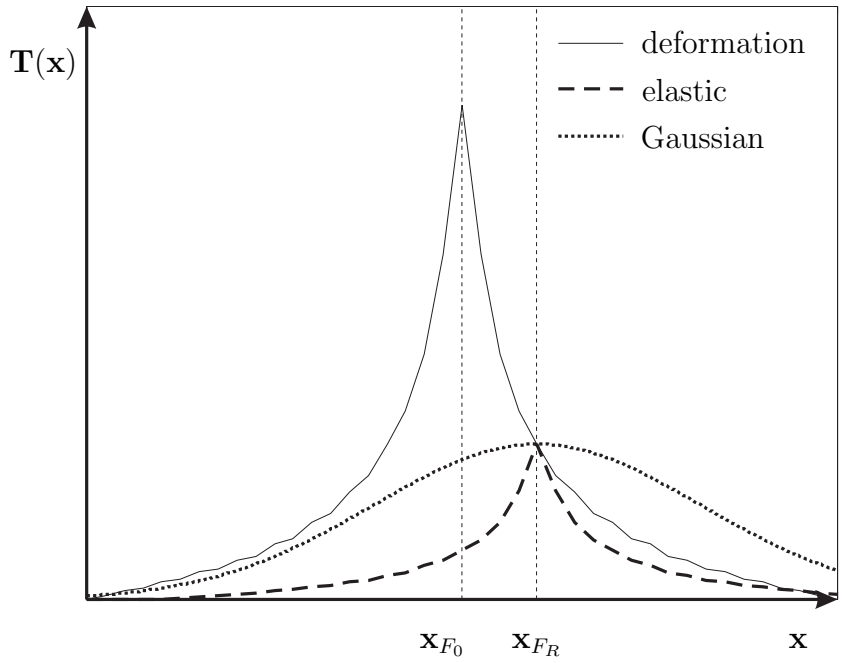


Figure 4.11: *Recovering of deformation of elastic material, deformed by the force F_0 that acts on the point inside the homogeneous image region (the deformation is presented by the thin solid line). Although the registration may always be able to make the images look the same, it cannot perfectly recover such deformation, because the forces F_R that are estimated during the registration act only on the edges of anatomical structures. In this case the convolution kernel that suits to the properties of the deformed media (dashed line), may cause larger registration error than some other less realistic kernels, e.g. Gaussian (dotted line). In the presented experiment the elastic kernel (dashed line) results in 30% larger RMS registration error than the Gaussian kernel (dotted line).*

sian kernel is used, it may produce more correct results. For example, in our experiment the optimal Gaussian kernel, see Figure 4.11 dotted line, performed 30% better than the elastic kernel, with respect to the RMS registration error e_{RMS} .

The experiments show that the best registration results are not necessarily obtained by spatial deformation models that exactly follow deformable properties of the anatomy. This justifies the use of Gaussian models, which are often used due to their lower computational cost. The selection and setting of the model is therefore not straightforward an depend on the application, specifically on the expected distribution of body forces, required compressibility and volume preservation, etc.

4.3 Discussion and conclusion

In this chapter we focused on spatial deformation models based on convolution. In the first part we devised the combined elastic-incremental model and compared it with the elastic model and the incremental model. The combined model tends to reduce the systematic error of the elastic model and the accumulated error of external forces, which is characteristic for the incremental model. It turns out that these two errors are related, such that the decrease of the first one increases the second one, and vice versa. The combined model enables to find an optimum, where the total error is the lowest. The comparison results prove that the combined model does perform better than the other two models individually. However, the optimal settings, including the ratio between elastic and incremental regularization, may depend on the nature of the deformation.

In the second part of the chapter we analyzed and compared two different convolution kernels: the elastic kernel and the Gaussian kernel. The elastic kernel models the real tissue properties better than the Gaussian kernel. Its major advantage is to provide control over the material compressibility and thus enable volume preservation. As expected it turned out that it gives better results when the anatomy is deformed by forces that act on the edges of anatomical structures. If this is not the case, the situation is more difficult and transformation cannot be recovered exactly. It turns out that in the case, when forces act in homogeneous image regions, the realistic elastic model gives even worse results than some other, nonrealistic models, because estimated external forces differ from the ones that actually deformed the tissues. We have illustrated this with the 1D experiment, in which the Gaussian kernel resulted in 30% lower RMS registration error than the elastic kernel. This justifies the use of Gaussian models, which furthermore have an additional advantage of lower computational cost.

The presented results indicate that the optimal convolution based spatial deformation model is the combined elastic-incremental model. Because in reality

the forces that deform the anatomies are not expected to be limited to the edges of anatomical structures, the elastic kernel is not necessarily the best one and computationally less demanding Gaussian kernel may be used instead. Because the best model does not necessarily suit to the realistic deformation properties, the setting of such model is difficult and depends on the application. It does not depend only on the properties of the real tissues, but also also on the expected distribution of body forces, which cause the deformation. In reality, it is difficult to take into consideration all the factors that have influence on the optimal setting of the spatial deformation model. Consequently, we expect that the best setting can be obtained experimentally.

5. Evaluation of Multi-Modality Non-Rigid Registration

Image registration is used to find a transformation, which maps one image to the other one, such that anatomical features are brought into a spatial correspondence. The evaluation of registration should therefore judge the correctness of the obtained transformation. In practice this is difficult, because the correct transformation is in reality never known, and consequently, evaluation must be accomplished using other approaches, based on synthetic deformation, image segmentation or direct image comparison. In this chapter we analyze these approaches and expose their limitations. Because none of these methods can reliably deal with all kinds of misregistration, we propose a three step evaluation strategy, which alleviates the problem by separately evaluating individual functionally independent stages of the registration.

5.1 Introduction

In the last decades, there has been a growing interest in developing new and better techniques for non-rigid registration, particularly in medical imaging field. A variety of non-rigid registration methods exist [82], and several validation schemes have been proposed [89], particularly on rigid registration [88]. In spite of this, objective evaluation of non-rigid image registration remains an open question. Following the definition of image registration, evaluation should judge the quality of a given registration according to the correctness of *transformation*, which for every point in the source image specifies the displacement with respect to its initial position. True transformation of real medical images, however, is generally not available, rendering such direct approaches difficult or even impossible. Other, indirect evaluation approaches, based either on synthetic deformation, manual and automatic image segmentation, or straight image comparison have been used instead [32, 36, 13, 41]. The danger with these indirect approaches is that they can lead to questionable interpretation of evaluation results. Let us imagine an image of a square, rotate this image by 30 degrees and then register it back to the original one using a non-rigid registration technique, see Figure 5.1. According to the Figure 5.1 the obtained image appears to be quite similar to the

original one, suggesting successful registration. Nevertheless, a detailed examination of the corresponding transformation shows that back-transformed image was not only rotated as required, but it was also largely deformed. Is this registration correct, as indicated by the image appearance, or is it in error due to the induced deformation? One could argue that, if the goal were to align the images region-wise, the result could be considered adequate, regardless of the actual mapping. But should the answer to this question depend only on the circumstances under which the method is to be used? Our next example involves two morphologically different images, see Figure 5.2. Depending on the deformation model non-rigid registration of such images can yield a number of different results. Which registration is better and which one is the correct one? Should registration make the resulting image as similar to the target as possible, or should it preserve morphologically important features of source image, although leaving much greater difference between the resulting and the target image? Thus, how to evaluate non-rigid registration results? We believe that these questions should be answered by following the definition of image registration.

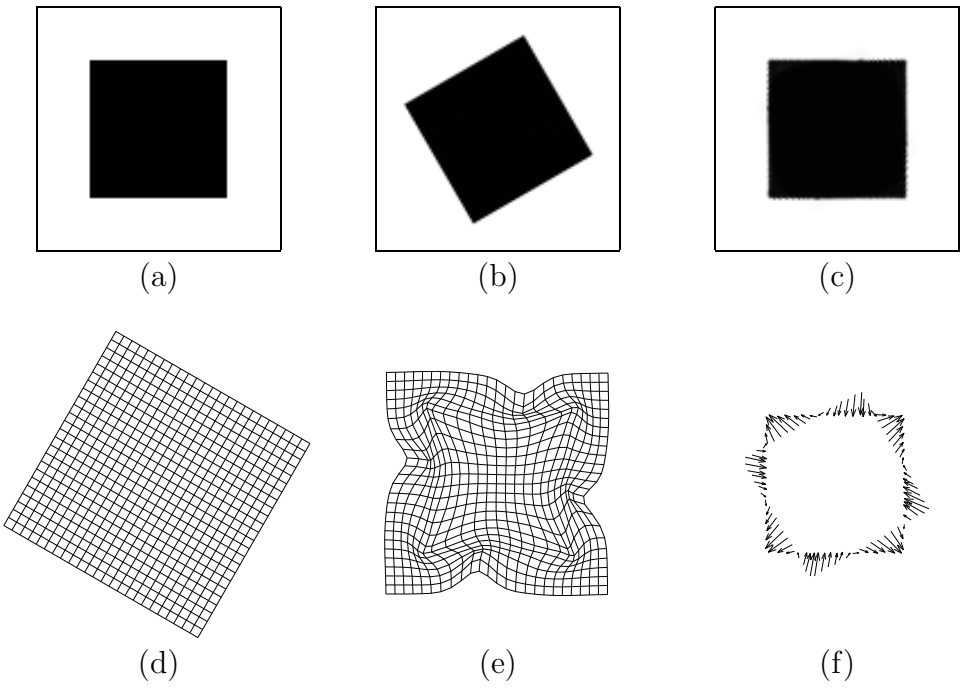


Figure 5.1: *Registration of squares, rotated by 30 degrees. (a) target image, (b) source image, (c) registration result, (d) and (e) expected and obtained transformation of image grid respectively, (f) obtained displacements shown for edge points.*

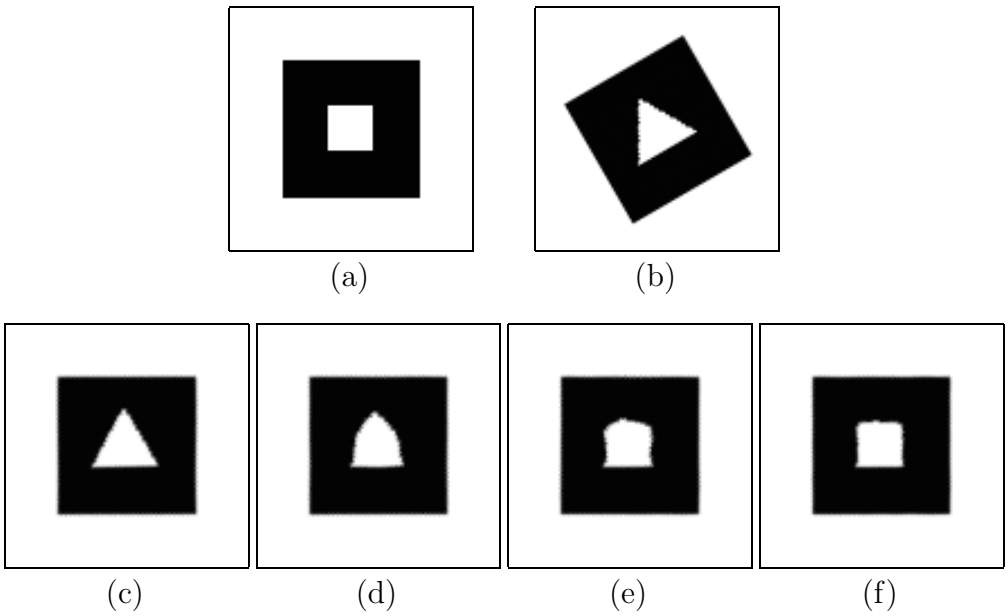


Figure 5.2: *Registration of images that have some morphological differences. (a) target image, (b) source image, (c-f) possible registered images.*

5.2 Non-rigid registration issues

Non-rigid image registration is in medicine used to bring homologous points in images that represent related anatomies or/and their functionalities into correspondence. In other words, the aim of registration is to transform one of the anatomy representations to match the other in a way of putting anatomically important features that are present in both images into correspondence. Registration therefore tries to geometrically transform one of the images, with the purpose of making the images more similar, but the basic problem associated with this is that the images can become similar even if anatomical features do not correctly correspond. This is clearly demonstrated in Figure 5.1. Although the transformed image seems to be aligned with the target, one can observe that pixels representing square corners are actually mapped to points on the sides (Figure 5.1f). Other points are mapped incorrectly as well. In this example the source of misregistration lies in inappropriate spatial model. Generally, the selection of spatial model should be based on the properties of modeled object or on the expected differences between target and source object/anatomy. In case of intrasubject registration, where both anatomies are approximately the same, spatial model should be based on deformable properties of tissues. Unfortunately real tissue properties are usually not exactly known, and furthermore, they differ according to tissue types. Thus, deformation model of registered anatomy is usually simplified in a way of using only one model for all the tissues, see Chap-

ter 4. The model should favor those kinds of deformation that are admissible (realistic) and penalize the deformations that are not. If a registration model does not enable correct deformation, e.g. if the model is too stiff, detailed differences remain and such registration can be characterized as underregistration. On the opposite, too loose models introduce new detailed differences (as shown in our example in Figure 5.1) and can be considered as overregistration.

Intersubject registration, due to high intersubject biological variability, makes the problem even harder. For example, each brain has a unique configuration of gyri and sulci. In fact, some sulci are not found in all individuals [57]. This causes difficulties when trying to precisely register such images. Nevertheless, precise registration may not be even desirable [89] and can be treated as overregistration. Although overregistered images may look more similar, they can not provide reliable conclusions about the morphological correctness, if the morphology can be arbitrarily changed by the registration. This phenomenon is illustrated in Figure 5.2. If angles are important for further studies, the solution that removes one angle and adds two new ones is inappropriate. Furthermore, would it be better to match the remaining angle of triangle with one of two remaining angles of square? Similar problems and questions can appear when matching sulci and gyri in intersubject brain studies. Is it therefore better to remove only those differences, which can be removed with reasonable level of reliability and accuracy [89]? Or maybe all the differences shall be removed as assumed in [35]? Visual appearance, or direct comparison of registered images is clearly not sufficient criterion for evaluation of image registration, as even visually similar images do not give adequate information about the correctness of deformation.

It is obvious that registered images will look more similar if a more loose model is used. By reducing the model stiffness, it may be possible to completely remove all visual differences (e.g. deform circle to C shape). The question remains whether such registration is also better than one obtained by using a more stiff model that may yield visually less appealing results? Registration should not tend to remove all image differences but only those that are anatomically plausible. Unfortunately it is not clear how to distinguish between plausible and implausible deformations. Perhaps the best we can do is to calibrate the registration algorithm, e.g. by setting stiffness of the spatial deformation model according to the application specific knowledge. Evaluations should take that knowledge into consideration as well.

5.3 Evaluation methods

Following the definition of image registration, evaluation should judge the registration by measuring remaining deviation of anatomical feature positions. Specifically, evaluation should measure the difference between obtained and correct

displacements for all image points that represent studied anatomy. To perform such evaluation the correct transformation must be known, but unfortunately this is not the case when registering real medical images. In general, there are two ways how to tackle this problem. Firstly, by using synthetic deformations, and secondly by using anatomical features that can be identified in both images. However, sometimes none of previous two approaches is possible, and the only method that remains applicable is a direct image comparison. In this section we analyze these three evaluation methods and identify their limitations.

5.3.1 Recovering synthetic deformations

In practice, the correct registration is available only if deformation is generated synthetically. In that case, registration can be evaluated as illustrated in Figure 5.3. The procedure requires two images, A and B, that are correctly registered, e.g. acquired at the same time, simulated or taken into correspondence by some validated registration approach. It is also possible to use the same image for A and B, but results may be biased because of correlated image noise or high correlation between image artifacts. Image A is then deformed using a known transformation $\mathbf{T}_0$ to image $\mathbf{T}_0 A$ which is then used as a target for registering image B. The result of registrering image B to image $\mathbf{T}_0 A$ is transformation $\mathbf{T}_R$, which should in the ideal case equal the synthetic transformation $\mathbf{T}_0$, but in reality it differs. The quality of registration can be measured by comparing the transformations. Such comparison can be based on measuring a RMS residual difference e_{RMS} between displacements $\mathbf{T}_0(\mathbf{x})$ and $\mathbf{T}_R(\mathbf{x})$:

$$e_{RMS} = \sqrt{\frac{1}{N_\Omega} \sum_{\mathbf{x} \in \Omega} \left(\mathbf{T}_0(\mathbf{x}) - \mathbf{T}_R(\mathbf{x})\right)^2}. \tag{5.1}$$

Voxels representing background do not provide reliable information about registration quality and may be removed by mask Ω. N_Ω denotes number of voxels accepted by the mask Ω and $\mathbf{x}$ are spatial coordinates. Similarly, it is also possible to compute mean, median and maximal registration errors.

The limitation of the evaluation based on synthetic deformations is in the synthetic nature of the transformation $\mathbf{T}_0$. If the transformation is not realistic, the obtained evaluation results may be biased, because spatial deformation models may systematically favor the simulated deformation. In order to perform unbiased evaluation of registration, the synthetic deformation must be generated by precise modeling of the deformed anatomy, such as in [75]. Evaluation by recovering synthetic deformations is therefore less suitable for intersubject registration, where deformations are difficult to be modeled due to high intersubject anatomical variability.

When the system for non-rigid registration consists of functionally independent stages, the evaluation based on synthetic deformations enable unbiased anal-

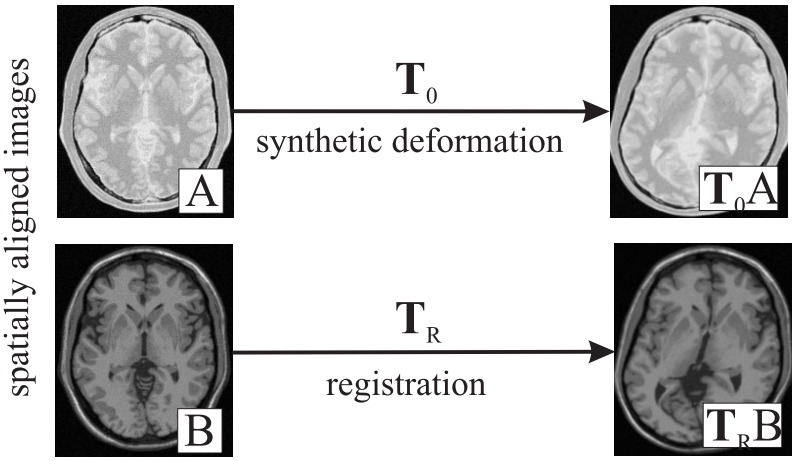


Figure 5.3: *Evaluation based on recovering synthetic deformations. Image A is synthetically deformed with transformation* $\mathbf{T}_0$*. The obtained image* $\mathbf{T}_0 A$ *is used as a target for registering image B. The evaluation is then based on comparison of transformations* $\mathbf{T}_0$ *and* $\mathbf{T}_R$*.*

ysis of performance for those stages that do not use spatial information. In the case of non-rigid registration based on point similarity measures, see Chapter 6., the only stage that uses spatial information is the spatial deformation model. The performance of the other stage, estimation of external forces, can thus be analyzed properly even by using less realistic synthetic deformations. Although the registration results depend on the capability of the spatial deformation model to fit to the synthetic deformation, the relative ranking of results reflects the suitability of the methods for estimation of external forces.

To illustrate the evaluation based on recovering synthetic deformations we compared different point similarity measures. Point similarity measures do not use any spatial information, and consequently, they can be compared using this approach. The experiment was performed using several different synthetic deformations, all of them generated as a sum of Gaussian functions. The difference between the deformations were in the amplitude of the Gaussian functions, $A = \{5, 10, 15\}$, and their standard deviation, $\sigma = \{10, 20, 30, 40, 50\}$. The obtained results are tabulated in Table 5.1. Relative performance of different point similarity measures is compared for different synthetic deformations in Figure 5.4.

The results show that similar conclusions about the performance of different point similarity measures can be made almost regardless to the synthetic deformation. The only limitation is that comparison cannot be made well when the spatial deformation model cannot fit to the synthetic deformation. In this case the obtained results are poor for all of the similarity measures, e.g. for $\sigma = 10$. In all the other cases, it can be concluded that measures S_P and S_U are not a good selection, other measures are better and the best measure is the measure

Table 5.1: *Recovering of synthetic deformations used for comparison of point similarity measures. Synthetic deformations were generated as a sum of Gaussian functions, using three different amplitudes, $A = \{5, 10, 15\}$ mm, and five different standard deviations, $\sigma = \{10, 20, 30, 40, 50\}$ mm.*

	$\sigma = 10$		$\sigma = 20$		$\sigma = 30$		$\sigma = 40$		$\sigma = 50$	
S	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}	e_{RMS}	e_{max}
$A = 5$										
–	0.30	6.41	0.87	6.23	1.67	6.04	2.70	6.75	3.91	7.89
S_P	0.42	7.08	0.94	7.29	1.70	7.39	2.38	7.77	2.85	8.25
S_H	0.38	4.85	0.56	4.09	0.78	3.75	0.85	3.27	0.82	3.32
S_{MI}	0.29	4.79	0.49	4.43	0.70	4.10	0.75	3.80	0.75	3.64
S_{PC}	0.26	5.13	0.46	3.32	0.68	3.12	0.71	2.90	0.69	2.50
S_{HC}	0.44	4.61	0.61	4.31	0.82	4.07	0.90	3.52	0.89	3.48
S_S	0.32	4.53	0.45	2.83	0.67	2.76	0.71	2.92	0.66	2.54
S_U	0.27	6.39	0.68	6.60	1.18	6.41	1.49	6.16	1.60	6.43
S_{UH}	0.32	4.47	0.55	4.29	0.74	4.09	0.84	3.79	0.79	3.49
$A = 10$										
–	0.60	12.82	1.74	12.46	3.34	12.08	5.40	13.51	7.83	15.79
S_P	0.66	13.54	1.75	13.59	3.14	13.04	4.21	13.42	5.29	13.92
S_H	0.51	11.62	0.89	9.06	1.01	8.75	0.84	4.53	0.83	3.88
S_{MI}	0.48	11.30	0.80	9.41	0.95	8.89	0.80	5.32	0.76	3.71
S_{PC}	0.50	11.76	0.86	10.45	0.93	7.61	0.81	5.46	0.79	4.68
S_{HC}	0.54	11.48	0.93	9.61	1.04	8.36	0.89	4.07	0.89	3.62
S_S	0.43	11.14	0.66	8.61	0.75	5.08	0.72	3.83	0.69	3.20
S_U	0.55	12.99	1.45	12.96	2.44	11.93	3.04	11.64	3.39	10.78
S_{UH}	0.49	11.23	0.80	9.51	0.97	9.21	0.81	4.42	0.80	3.74
$A = 15$										
–	0.89	19.23	2.61	18.69	5.02	18.12	8.11	20.26	11.74	23.68
S_P	0.94	20.00	2.59	19.89	4.59	18.67	6.74	20.10	6.70	17.57
S_H	0.78	17.91	1.51	14.71	1.57	10.92	1.03	7.79	1.02	6.49
S_{MI}	0.75	17.55	1.48	14.23	1.61	12.16	1.19	9.38	0.94	6.10
S_{PC}	0.75	18.05	1.56	17.18	1.86	13.74	1.91	11.55	1.74	10.15
S_{HC}	0.78	17.99	1.56	14.60	1.47	9.95	1.10	7.91	1.08	6.60
S_S	0.70	17.65	1.28	14.95	1.11	9.61	1.00	8.15	0.75	4.68
S_U	0.85	19.58	2.30	19.40	3.87	17.59	5.42	18.05	5.52	16.02
S_{UH}	0.75	17.65	1.46	14.30	1.53	11.85	1.03	8.06	0.89	5.76

All results are in millimeters and represent residual errors after the registration.

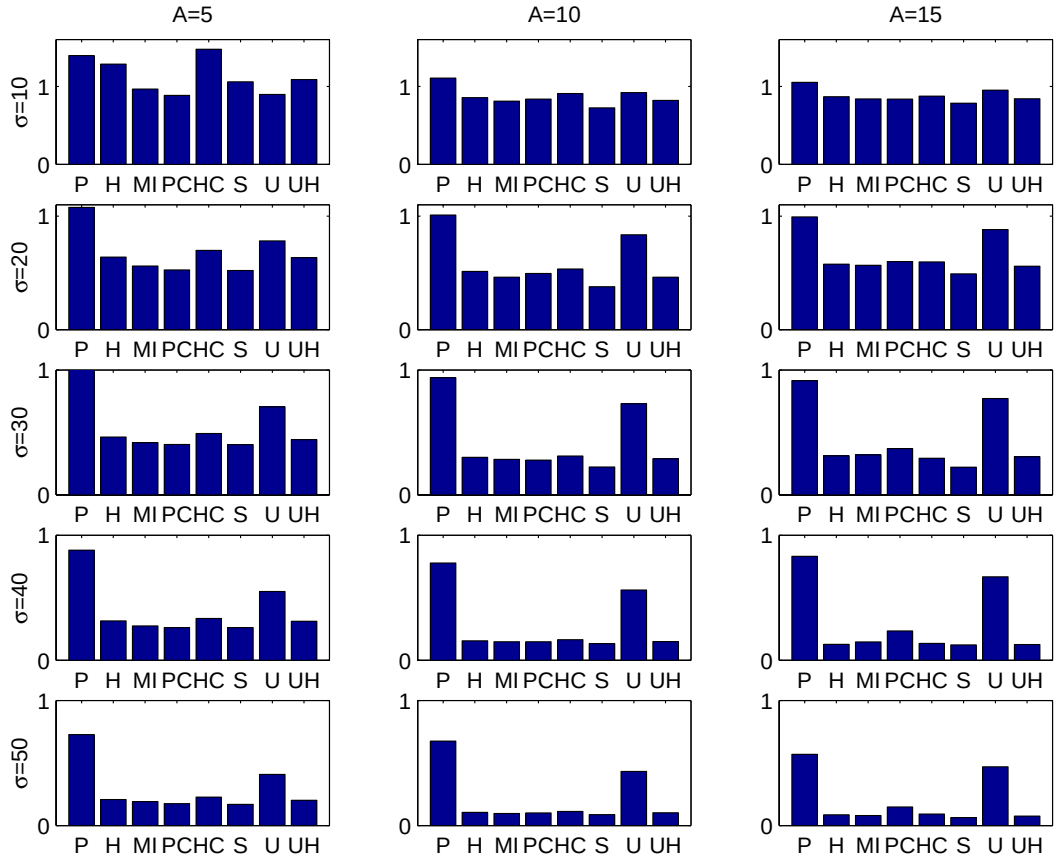


Figure 5.4: *Comparison of point similarity measures based on recovering synthetic deformations. The synthetic deformations were generated as a sum of Gaussian functions, using three different amplitudes,* $A = \{5, 10, 15\}$ *mm, and five different standard deviations,* $\sigma = \{10, 20, 30, 40, 50\}$ *mm. Graphs show residual RMS errors* e_{RMS} *relatively, according to the initial error.*

S_S. Of course, the results could differ if using different type of data (in our case MRI-PD and MRI-T1 images were used). Because the same conclusion can be made for very different deformations (different σ), this proves the functional independence of point similarity measures and the spatial deformation model, and furthermore, this also proves the suitability of evaluation based on recovering synthetic deformations for such comparisons.

5.3.2 Segmentation based evaluation

The purpose of image registration is to put anatomical features in both images into correspondence. Registration can therefore be evaluated by measuring the correspondence of anatomical features. For that purpose images must be segmented to know the positions of features in both images. Different types of features can be used, e.g. point landmarks, segmented structures or segmented tissue types. The problem of using point landmarks is in finding the correct homology, as even experts are likely to disagree about the correct position of landmarks in some brain regions [89]. Extraction of three-dimensional anatomical structures is much easier, although it is still tedious. Correspondence of segmented regions can be measured using the same methods as used for evaluation of image segmentation [30]. One of the most widely used overlap measures is a Jaccard coefficient J, defined as intersection of corresponding regions on both images (R_A and R_B) divided by their union:

$$J = \frac{R_A \cap R_B}{R_A \cup R_B} \tag{5.2}$$

where $J = 1$ means perfect agreement and $J = 0$ complete disagreement. However, evaluation based on segmented structures can only judge position of region surfaces and cannot detect incorrect matching within individual structures, which can be caused by overregistration. For example, the evaluation of registration shown in Figure 5.1 would not detect any irregularity. Furthermore, it is not possible to distinguish between registration errors and true morphological variability. For example in Figure 5.2 evaluation would show best overlap for registration result Figure 5.2f, which is morphologically different and therefore possibly incorrect. Instead of using segmented structures evaluation can also be based on segmented tissues. This approach is even more problematic, as all previously mentioned problems remain, but segmented regions are larger and consist of numerous structures, which are not distinguished. On the other hand segmentation of tissues is much easier than segmentation of structures and can be performed automatically.

One of important issues concerning evaluation based on segmentation is quality of the segmentation. Evaluation cannot be more reliable than segmentation itself, as evaluation only estimates the overall error that comprises registration as

well as segmentation incorrectness. Note that registration can morphologically change the source image, so images must be segmented prior the registration.

An additional difficulty of segmentation based evaluation is the need that the same anatomical features can be identified in both of the images. This is problematic in the case of multi-modality registration, where different images represent different physical material properties, and as such, only a limited (small) number of corresponding anatomical features can be segmented in both of the images.

For the illustration we show how segnentation based approach can be used for evaluation of intersubject registration of MRI brain images. Segmentation of certain brain structures is a tedious task and such images are rare and valuable. As our dataset only contains some 2D segmented regions, measuring of 3D overlap is not applicable. However, visualization of such curves overlayed on registered image can be used for evaluation performed by medical experts (see Figure 5.5).

We have also evaluated intersubject registration using segmentation to different tissue types. We have obtained 60% overlap of white matter and 54% of gray matter. However, as mentioned, such results depend on the quality of segmentation and furthermore, they do not distinguish between registration error and real morphological differences.

5.3.3 Evaluation based on registered images

If segmentation is not available for both images (target and source), evaluation can only rely on images themselves. It should be used with extreme caution as images can become similar even because of incorrect registration or overregistration. Still, for images that are not overregistered and are obtained by using a sufficiently stiff model, better registration in general also reflects in more similar images. Evaluation based on registered images can be made either by measuring global image similarity or by providing visual comparison, without numerical evaluation. Results can be visually shown as an overlay of one image on top of the other, either by using different methods for showing the whole images (image difference, chessboard image), or only by showing important edges or points. Results presented by using this approach often require some expert knowledge, as results are actually not yet evaluated. Such results can be furthermore numerically evaluated by experienced users, e.g. medical experts, specially if specific knowledge is required (e.g. for identifying certain structures on both images).

In some cases evaluation based on resulting images seems to be unavoidable. One of the examples is PET-CT registration of thoracic images. Although this is intrasubject registration there are large differences between the images due to different respiration during image acquisitions. An example of such registration results obtained by our non-rigid registration approach is shown in Figure 5.6.

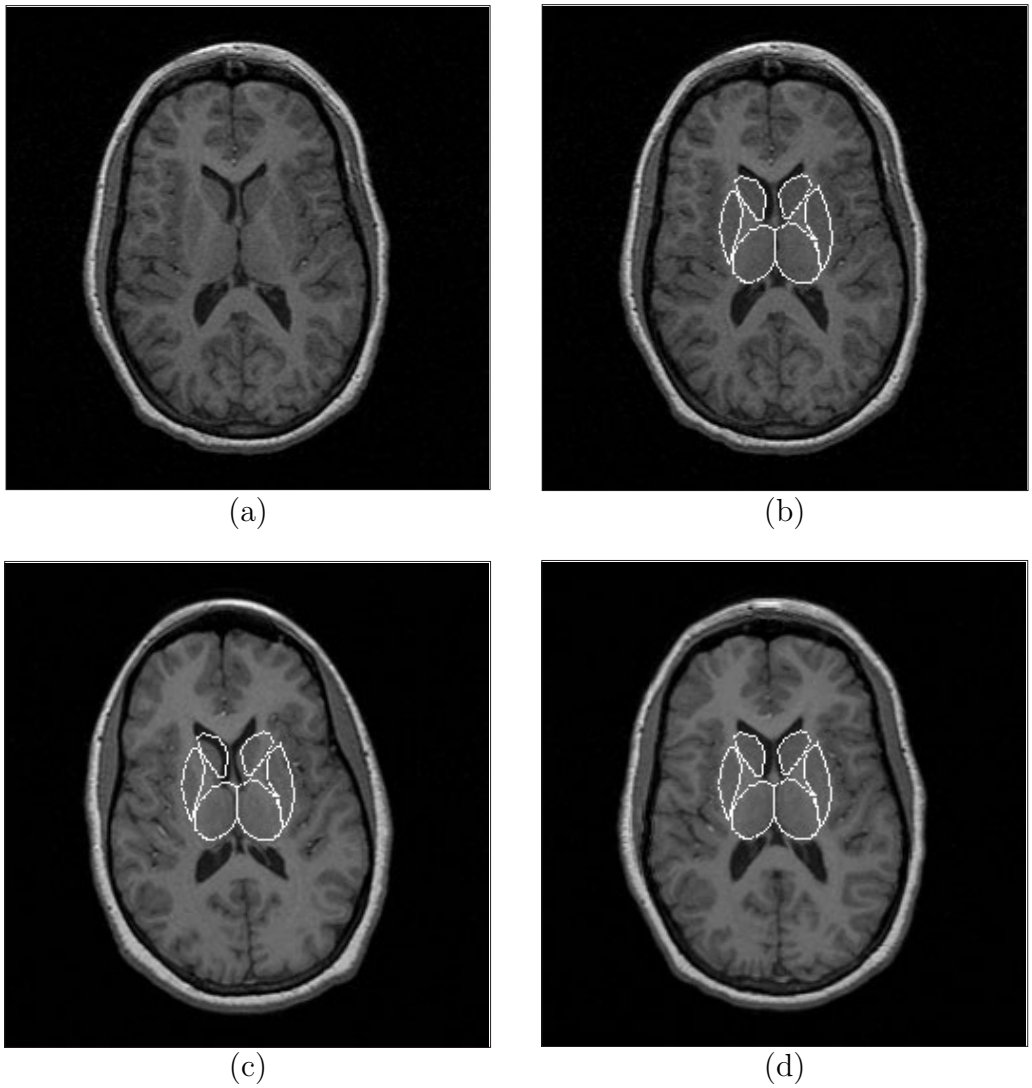


Figure 5.5: *Intersubject registration of MRI head images. Target (a), regions of interest segmented on the target image (b), segmented regions overlayed on rigidly and nonrigidly registered source image (c,d).*

5.4 Three step evaluation strategy

Evaluation of the registration systems is difficult, because all the evaluation methods presented earlier have some limitations, which can result in evaluation errors. The recovering of synthetic deformations may be biassed due to unrealistic synthetic deformations, while the other methods rely only on matching of edges of anatomical structures and thus cannot assess all the transformation components, which renders detection of overregistration impossible. Furthermore, there is ad-

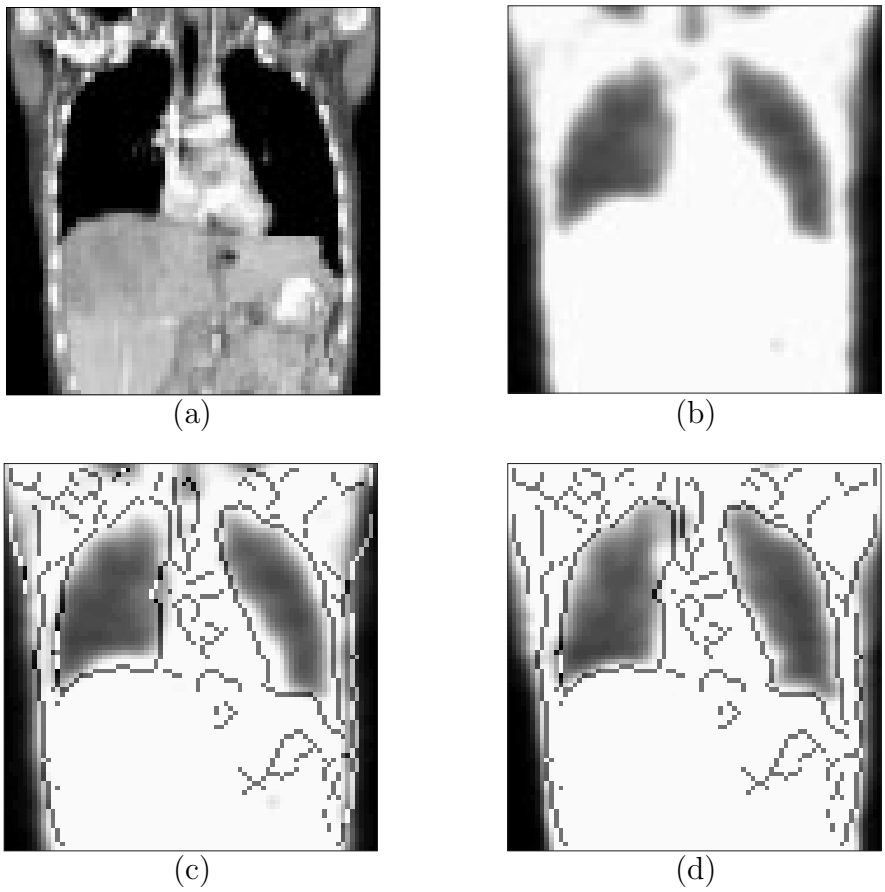


Figure 5.6: *PET-CT registration of thoracic images: CT image (a), PET transmition image (b), rigidly registered (c) and nonrigidly registered (d) PET transmition image overlayed with contours obtained using Canny edge detector on CT image.*

ditional difficulty when evaluating multi-modality registration: the images may look very different and only a small number of anatomical features may be identified in both of them. To alleviate the problem of evaluation of multi-modality registration systems, we propose a three step evaluation strategy.

When the system for multi-modality non-rigid registration can be separated into functionally independent stages, this stages can be evaluated separately and independently of the whole system. Non-rigid registration can sometimes be separated into two functionally independent stages: the external stage, which drives the registration, and the internal stage, which regularizes the registration. In the case of our registration system described in Chapter 6. the external stage is the estimation of external forces, and the internal stage is the spatial deformation model. It is important, that the external stage does not provide any regulariza-

tion, because it drives the registration on the basis of individual points (thanks to the point similarity measures used, see Chapter 2.).

We have already shown that the external stage can be evaluated independently to the internal stage, by recovering synthetic deformations. This avoids the problem of generally unknown realistic deformations. Such evaluation can be used to compare and select methods of this stage (e.g. similarity measure, force estimation approach, etc.) and set them appropriately. For this purpose realistic images should be used, while the deformations do not necessarily need to be realistic and can be generated synthetically.

The internal stage (spatial regularization) can be evaluated separately as well. It requires evaluation based on realistic deformations, but however, it does not depend on actual images, which thus do not need to be of different modality. The segmentation based evaluation can be used, where the actual transformation can be assessed from mono-modality images instead of multi-modality ones. This alleviates the problem that in general only a small number of anatomical features can be identified in both of the images when they are of different modality. Consequently, this approach enables more precise setting of the internal stage than evaluation of the whole system.

Even if both of the stages are evaluated separately, this does not provide an absolute evaluation of the whole system. For this purpose, the whole system needs to be evaluated using realistic images as well as realistic deformations. Furthermore, segmentation based evaluation can be used if the segmentation is available, or it can be based only on the registered images if this is not the case. However, if both of the stages are already evaluated and appropriately set, they do not need to be set again. Consequently, evaluation of the whole system requires lower number of images, and furthermore, it reduces the possibility for overregistration.

To summarize, the three step evaluation strategy consists of the following three steps:

1. Evaluation and settings of external stage based on realistic images and recovering of synthetic deformations. The synthetic deformations do not need to be realistic.

2. Evalaution and settings of internal stage based on realistic deformations and segmentation based evaluation. The images do not need to be of different modality.

3. Evaluation of the whole system using realistic images and realistic deformations, and segmentation based or image based evaluation.

5.5 Conclusion

In this chapter we have illustrated some of the problems pertaining to evaluation, and discussed their origins. Because of the unknown transformation that would ideally register real medical images, evaluation must be accomplished using other approaches, based on synthetic deformation, image segmentation or direct image comparison. However, all these approaches can lead to certain evaluation error, as no method can reliably deal with all kinds of misregistration.

Evaluation based on synthetic deformation is less appropriate for evaluation of whole registration systems, because it favors deformations that are similar to the synthetic ones. It is still appropriate for comparison of methods that are not related to the spatial model as well as for testing overall registration capabilities. Segmentation based evaluation cannot detect misregistration within segmented features and furthermore does not distinguish between registration errors and true variability. However, by using a large number of correctly segmented smaller features, such evaluation converges towards the ideal evaluation. Evaluation based only on the images is the least reliable as images can also become similar due to inappropriate morphological changes. To improve the evaluation we have proposed a three step evaluation approach, which is based on independent evaluation of two registration steps: estimation of external forces and spatial deformation model.

One of the non-rigid registration problems is overregistration. It often cannot be detected even by using segmentation based evaluation. This problem is related to settings of spatial deformation model. If a more loose model is used, resulting images may be more similar even if deformation is worse. On the other hand, too stiff models lead to underregistration that prevents correspondent structures to be registered correctly. Therefore, a spatial model should be tuned to the problem under consideration.

There is also a question of registering morphologically different structures. Should morphological differences remain or not? Most evaluation methods cannot distinguish between registration errors and true variability, so "good"registration results can be obtained only by changing source image morphology. If morphological differences are supposed to be preserved, registration can be correctly evaluated only when it is controlled by medical experts.

The results shown are based on the evaluation of our non-rigid multi-modality registration method presented in Chapter 6. However, the results shown in this chapter only serve for the illustration of different evaluation methods. Evaluation of registration components, i.e. similarity measures, force estimation methods and spatial deformation model is provided in Chapters 2., 3. and 4. respectively.

6. Image Registration System

In this chapter we present an image registration system, which we have designed to support the development of approaches presented in previous chapters of this thesis. The system is based on point similarity measures and can be used for multi-modality rigid and non-rigid registration. The system was not built with some specific application in mind, but to demonstrate general solutions, which could be used for all kinds of 3D medical image registration problems. As such it does not incorporate any prior knowledge specific to certain modality or anatomical structures. Due to the generality of the system it can be used to demonstrate the capabilities of medical image registration and to aid to the future developments and medical research. The system was also used to perform the experiments shown in this thesis.

6.1 Introduction

We have built an image registration system to enable development of the proposed methods and their testing. The system is based on the methods presented in previous chapters. It employs point similarity measures, the symmetric registration approach and the combined elastic-incremental spatial deformation model. It does not employ any application specific knowledge such that it can deal with a variety of registration tasks: rigid and non-rigid, mono-modality and multi-modality, intra-subject and inter-subject, and for any part of the body. As such it can be used not only for evaluation of the proposed methods, but also for demonstration of capabilities of non-rigid registration and further medical research. It is not suitable for use in clinical medicine, as it was not designed for that purpose. For clinical tasks additional task-specific knowledge could be employed and the system should be extensively evaluated and validated.

6.2 Implementation

There are two types of registration implemented in the system: rigid and non-rigid. They represent two basic registration steps, which can be combined arbitrarily to obtain the final registration. Both, rigid and non-rigid registration

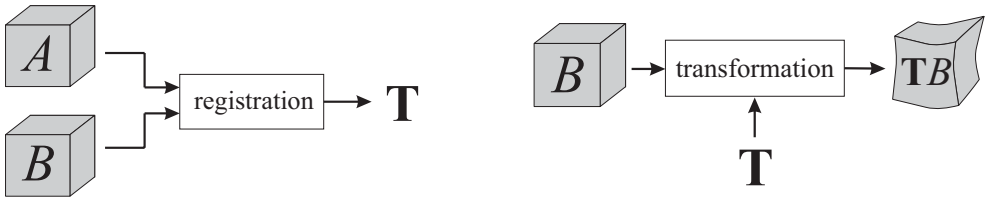


Figure 6.1: *The result of the registration is the transformation* $\mathbf{T}$ *(left). The registered image* $\mathbf{T}B$ *can be obtained after the registration by applying the transformation* $\mathbf{T}$ *to the source image* B *(right).*

steps allow multi-modality data. They are both based on point similarity measures, which adapt to the intensity dependence of the registered data, and do not use image processing algorithms that presume linear intensity dependence, including intensity interpolation and image resampling. Approaches that presume certain intensity dependence introduce new intensity values, which are originally not present in the images, and lead to unpredictable changes of intensity distributions, which may deteriorate the registration [50]. Because no image resampling is performed during the registration, the result of registration does not include the registered image $\mathbf{T}B$, but only the transformation $\mathbf{T}$, which puts the source image B into a spatial correspondence with the target image A. The registered image $\mathbf{T}B$ can be obtained after the registration by applying the transformation $\mathbf{T}$ to the source image B, see Figure 6.1. The registration may be performed in several resolution levels, such that each registration step may operate on arbitrary resolution level. This enables different multi-resolution strategies, which increase registration speed, avoid local extrema of similarity and thus improve the registration.

6.2.1 Rigid registration

Rigid registration is usually performed such that optimal parameters of rigid transformation are searched by some optimization procedure that maximizes a global similarity between the images. However, function of similarity with respect to transformation is not smooth, it includes local extrema, and furthermore, a global extremum may not appear exactly at the best image alignment. To overcome these problems several optimization procedures [51] and similarity measures [62, 39] have been introduced. Our implementation follows the same concept, but however, it does not use conventional similarity measures. Instead, the computation of global image similarity is based on point similarity measures, as described in Chapter 2.

A global image similarity S_G can be computed by averaging point similarities $S(v)$ over all image voxels v, or, as shown in Eq. (2.14), directly from joint inten-

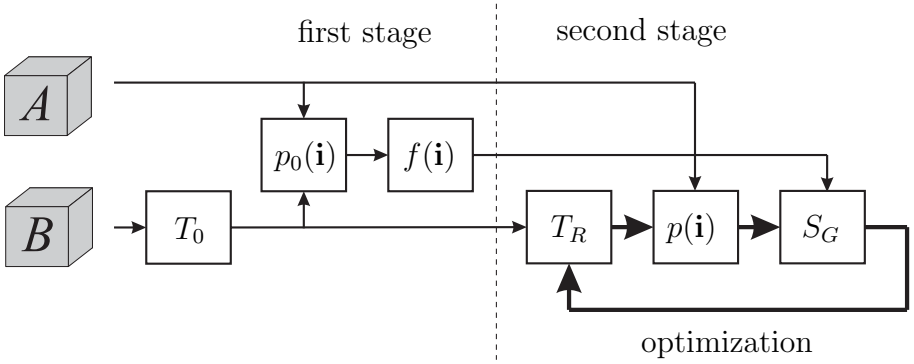


Figure 6.2: *A scheme of rigid registration based on point similarity measures. The registration consists of two stages: in the first stage the point similarity function $f(\mathbf{i})$ is estimated and in the second stage optimization is performed. The optimization loop is presented by thicker lines.*

sity probabilities $p(\mathbf{i})$ and point similarity function $f(\mathbf{i})$. Here, the point similarity function $f(\mathbf{i})$ needs to be computed only once, while the joint probabilities $p(\mathbf{i})$ must be reestimated for each different transformation $\mathbf{T}$. In accordance to this, the rigid registration consists of two stages, as illustrated in Figure 6.2. In the first stage the point similarity function $f(\mathbf{i})$ is computed from the target image A and the source image B transformed by an initial transformation $\mathbf{T}_0$. The initial transformation is provided by the previous registration step, or initially, it is set to identity $\mathbf{T}_0 = \mathbf{I}$ such that $\mathbf{T}_0(\mathbf{x}) = 0$. The point similarity function $f(\mathbf{i})$ can be obtained by any of the point similarity measures, from initial joint intensity distribution $p_0(\mathbf{i})$, which is estimated using partial volume interpolation [50]. In the second stage of the rigid registration, an additional rigid transformation $\mathbf{T}_R$ is searched by optimization. The criterion function is the global image similarity S_G:

$$S_G = \sum_{\mathbf{i}} p(\mathbf{i}) f(\mathbf{i}). \tag{6.1}$$

Here, point similarity function $f(\mathbf{i})$ is provided by the first registration stage, while joint intensity distribution $p(\mathbf{i})$ is reestimated for each different transformation $\mathbf{T}_R$ using partial volume interpolation. Measuring the global similarity in such a way is computationally efficient, as all complex operations that may be required to compute the point similarity function $f(\mathbf{i})$ are performed only once. In addition, this approach avoids interpolation artifacts, illustrated in Figure 2.6 where mutual information (MI) is compared to point based global similarity. Finally, this approach may also improve the position of global maximum of similarity, as in the example in Figure 2.8. The optimization is performed using Powell's method. The rigid transformation $\mathbf{T}_R$, which is searched by the optimization, is defined by six parameters: three translations (d_1, d_2 and d_3) and three rotations

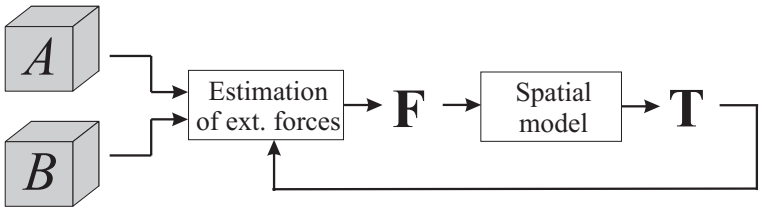


Figure 6.3: *A scheme of non-rigid registration.*

(roll ψ, pitch ϕ and yaw θ):

$$\mathbf{T}_R = \begin{bmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi\sin\psi - \sin\theta\cos\psi & \cos\theta\sin\phi\cos\psi + \sin\theta\sin\psi & d_1 \\ \sin\theta\cos\phi & \sin\theta\sin\phi\sin\psi + \cos\theta\cos\psi & \sin\theta\sin\phi\cos\psi - \cos\theta\sin\psi & d_2 \\ -\sin\phi & \cos\phi\sin\psi & \cos\phi\cos\psi & d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.2}$$

This transformation moves each point from original position $\mathbf{x}$ to a new position $\mathbf{x}'$ such that

$$\begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} = \mathbf{T}_R \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \tag{6.3}$$

Finally, the overall transformation $\mathbf{T}$, which is the result of the non-rigid registration, is obtained by applying the transformation $\mathbf{T}_R$ after the initial transformation $\mathbf{T}_0$:

$$\mathbf{T} = \mathbf{T}_R\mathbf{T}_0. \tag{6.4}$$

Because the sensitivity of point similarity measures depends on the quality of image match used for estimating the point similarity function $f(\mathbf{i})$ it is recommended that rigid registration consists of (at least) two rigid registration steps (in the same resolution level), as described in Section 6.2.3.

6.2.2 Non-Rigid registration

The non-rigid registration implemented in our registration system follows the continuum mechanic approach. A block scheme is shown in Figure 6.3. The registration consists of two functionally independent stages: estimation of external forces, which drive the registration in order to improve similarity of the images, and spatial deformation model, which spatially regularizes registration in order to suit the deformable properties of the registered anatomy. Both stages are iterated by the predefined number of iterations.

Estimation of external forces is implemented as described in Chapter 3. There are two methods implemented: a symmetric method and a conventional unidirectional method. It was shown that the symmetric method yields more consistent and also more correct registration results. Both methods are implemented using

point similarity measures. In contrast to rigid registration, a point similarity function $f(\mathbf{i})$ is in this case updated in each registration iteration to obtain optimal sensitivity. This is admissible because only similarities of point pairs obtained in the same iteration need to be compared. These similarities are always based on the same $f(\mathbf{i})$. External forces are computed for each individual voxel of image B independently. The position of voxel points in a world coordinate system $\mathbf{x}$ can easily be computed from their image coordinates $\mathbf{x}_B$ and transformation $\mathbf{T}$, and is updated in each iteration:

$$\mathbf{x} = \mathbf{x}_B + \mathbf{T}(\mathbf{x}_B). \tag{6.5}$$

Because external forces are computed for individual voxels with known position $\mathbf{x}$, there is no need for image resampling. Furthermore, because of point similarity measures, it is also possible to avoid interpolation of intensities, see Section 2.5.2. Consequently, presumption of (linear) intensity dependence can be completely avoided and registration can better adapt to the actual intensity dependence estimated from the images. Another advantage due to point similarity measures is that estimation of external forces does not influence the spatial regularization, because it is based on similarity of individual points rather than larger image regions. Consequently, spatial regularization remains only in the domain of the spatial deformation model in the second non-rigid registration stage. This makes setting and evaluation of the system easier, as each of the stages can be set and evaluated independently, see Section 5.4.

The spatial deformation model used in the system uses the combined elastic-incremental approach, which is described in Section 4.1.4. Here, the convolution kernels are Gaussian, due to their low computational cost. Standard deviations of the Gaussian kernels can be set arbitrarily to enable different model stiffnesses and different ratios between elastic and incremental regularization. The parameter k_F, which controls the registration speed, is chosen such that maximal possible change of transformation $\mathbf{T}(\mathbf{x})$ equals the size of one image voxel. The required number of iterations depends on the initial image mismatch. If a multi-resolution strategy is used, see Section 6.2.3, each registration step requires from 10 to 20 iterations.

6.2.3 Multiresolution strategy

In order to increase registration speed, avoid local minima and improve registration correctness, the overall registration can be performed in multiple resolution levels. Specifically, each registration step (rigid and non-rigid) can run in an arbitrary resolution level. Resolution levels are defined by resolution of the source image B, such that in resolution level 0 an original image B is used ($B^{(0)} = B$), while in higher resolution levels image resolution is lower. Specifically, image $B^{(k)}$ of the resolution level k is obtained by subsampling image $B^{(k-1)}$ of the resolution

level $k-1$ by the factor of 2. However, because images may have non-rectangular grid it may not be reasonable to subsample images in all spatial dimensions, such that subsampling is not performed in dimensions in which grid step already exceeds the minimal grid step (among all dimensions) by the factor of 2.

Note that discrete images have a limited frequency spectra and only frequency components that appear in both of the registered images can contribute to registration. Other frequency components may cause local extrema of similarity with respect to transformation $\mathbf{T}$. Consequently, it is important that image A is subsampled as well. An appropriate resolution of image A needs to be selected for each resolution level k, such that resolutions of both images, $A^{(k)}$ and $B^{(k)}$ are similar. Subsampling of image A is performed in the same way as subsampling of image B, the only difference is in number of subsamplings, which in this case does not necessarily equal the resolution level k. An optimal number of subsamplings n_k for resolution level k is defined as

$$n_k = \min\left(\left\lfloor \log_2 \frac{s(B^{(k)})}{s(A)} \right\rfloor, 0\right), \tag{6.6}$$

where $s(B^{(k)})$ and $s(A)$ denote minimal grid step (among all dimensions) of images $B(k)$ and A respectively.

The registration is performed in multiple steps, such that transformation obtained in one step is used as the initial transformation in the next step. Because every step can operate in arbitrary resolution level the transition between resolution levels is needed. In the case of rigid registration the transtion between two resolution levels does not require any modification of transformation, because rigid transformation is parametric and defined according to the world coordinate system. On the other hand, in the case of non-rigid registration, transition between resolution levels requires some processing, because the transformation is defined by displacements of image voxels. Whenever resolution of source image $B^{(k)}$ changes, the displacement filed needs to be resampled accordingly.

In order to prevent aliasing, each resampling of data requires filtering. For this purpose Gaussian filter is commonly used, but due to its disadvantage of producing new intensity values, which were originally not present in the images, it is less appropriate for filtering multi-modality images. Consequently, we use it only for subsampling non-rigid transformations, while for subsmpling the images we use $3 \times 3 \times 3$ median filtering. The advantage of median filtering is that it does not introduce new intensity values and cause minimal changes of intensity distribution. This enables usage of potential prior knowledge of intensity distributions in all resolution levels, irrespective to which image resolution the prior knowledge originally corresponds.

In general, registration steps can be arbitrarily threaded into the overall registration, such that each step can be rigid or non-rigid and performed in arbitrary

resolution level. However, note that higher resolution levels (lower image resolution) enable faster but coarser registration, with less local extrema of the similarity, and larger capture range, while lower resolution levels (higher image resolution) enable more precise registration. Furthermore, the quality of non-rigid image registration depends on the initial image misalignment, such that better registration results can be expected when the initial misalignment is lower. Due to all this we recommend that rigid registration is performed first, and after that the obtained transformation is improved by non-rigid registration. Both, rigid and non-rigid registration are recommended to follow the multi-resolution strategy shown in Figure 6.4. Here, the registration is performed in K resolutions ($k = 0..K-1$). It starts at the highest resolution level $k = K-1$ with the lowest image resolution, continues at lower registration levels and finishes at level 0, ($k = 0$), where the highest resolution images are used. Each resolution level may consist of multiple registration steps, such that resolution level k consists of L_k registration steps $l = 1..L_k$. Each of the registration steps can be set independently.

The registration starts from the initial transformation, which can be either set to identity, $\mathbf{T}_0 = \mathbf{I}$; $\mathbf{T}(\mathbf{x}) = 0$, or provided by previous registration, such as in the case of non-rigid registration, where the initial transformation is obtained by rigid registration. The initial transformation of the registration step $l+1$ is always set to the result of the previous registration step l:

$$\mathbf{T}_0^{(k,l+1)} = \mathbf{T}^{(k,l)} \quad ; \quad l = 1..(L-1). \tag{6.7}$$

or when changing the resolution level,

$$\mathbf{T}_0^{(k,1)} = \mathbf{T}^{(k+1,L)}. \tag{6.8}$$

For rigid registration it is recommended that multiple registration steps are performed in each resolution level. In practice we found that three registration steps, $L = 3$, in each resolution level are sufficient in most of the cases. Non-rigid registration does not require multiple registration steps in each resolution level such that we use a single registration step, $L = 1$.

6.3 Results

The system for multi-modality image registration was tested separately for rigid and non-rigid registration. Rigid registration was evaluated as a part of "The Retrospective Image Registration Evaluation Project"(RIRE), [88], which was designed to compare retrospective CT-MR and PET-MR registration techniques developed by different research groups. The results obtained by our system are presented in Section 6.3.1.

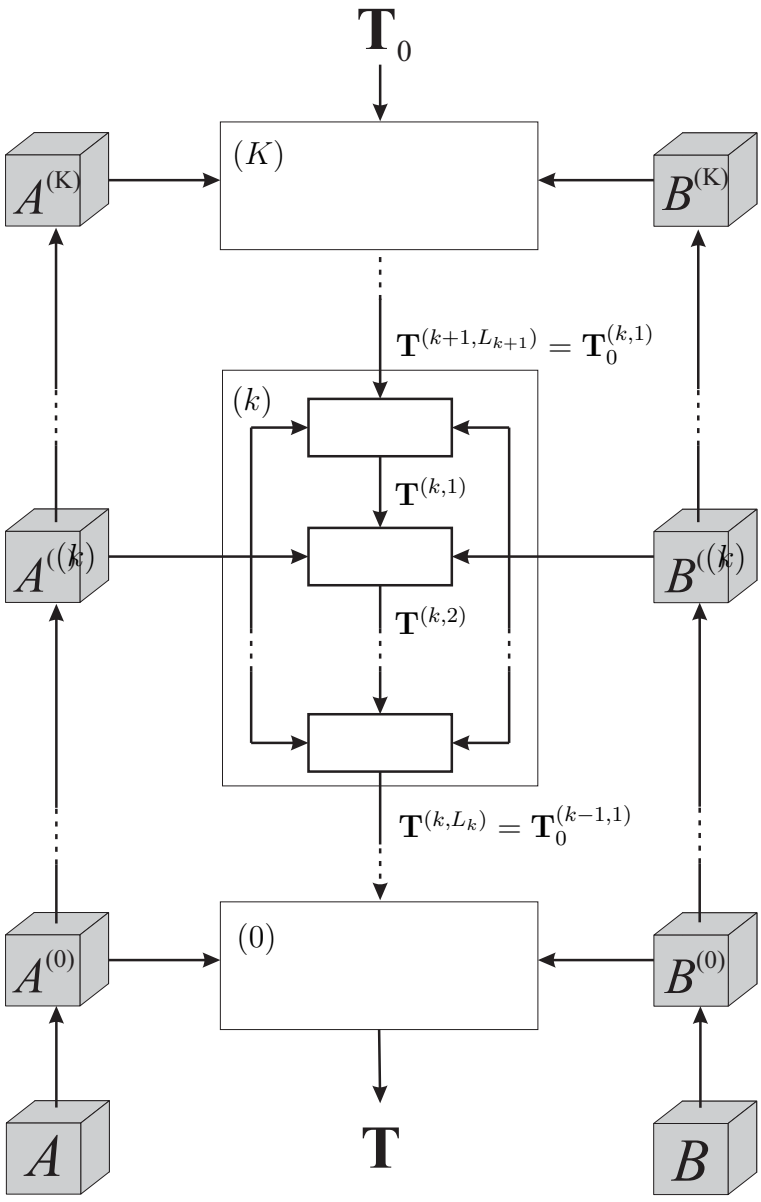


Figure 6.4: *Multi-resolution registration scheme, used for both, rigid and non-rigid registration. Registration is performed in K resolutions ($k = 0..K-1$), each of them consisting of L_k registration steps ($l = 1..L_k$). Each registration step can be set independently. The initial transformation $\mathbf{T}_0$ is either set to identity, $\mathbf{T}_0 = \mathbf{I}$; $\mathbf{T}(\mathbf{x}) = 0$, which is common for rigid registration, or provided by previous registration, as required by non-rigid registration.*

In contrast to the rigid registration, the non-rigid registration was not compared to other non-rigid registration techniques. Evaluation of non-rigid registration is more difficult than evaluation of rigid registration, see Chapter 5. The reason is in unknown non-rigid transformations in the case of registering real medical images, which cannot be assessed using prospective techniques, such as marker-based techniques can be used to assess rigid transformations. Furthermore, the evaluation of non-rigid registration methods is always application specific, because different applications may require (at least) different (settings) of spatial deformation model. Consequently, non-rigid registration systems cannot be directly compared such as rigid registration systems are compared by RIRE. Even if the comparison is made, it is indirect and application specific, such as [35].

Evaluation of individual components used in our system was presented in Chapters 2., 3. and 4. In this section we present a number of experiments, which have been performed within the scope of medical research, see Section 6.3.2.

6.3.1 Rigid registration

The rigid registration was evaluated by RIRE [88], (the project was formerly called RREP - "The Retrospective Registration Evaluation Project"), which was designed to compare retrospective CT-MRI and PET-MRI registration techniques used by a number of groups. It involves the use of FTP image database to allow the downloading of image volumes on which the registrations are to be performed. The idea is that the collaborating groups perform registrations on the image volumes, using their own retrospective techniques, and the group at Vanderbilt University evaluates the accuracy of these transformations by means of their own prospective, marker-based technique.

The image database includes images of 18 subjects, marked pt001-pt009 and pt101-pt109. The evaluation of registration accuracy was obtained by measuring registration errors for 10 points in brain anatomy marked as VOI1-VOI10. From all the results (for all points in all registered images) mean, median and maximum error is computed.

CT to MRI registration

Registration of CT images to MRI-T1 images was performed using 16 available CT-MRI image pairs (CT images for patients pt008 and pt009 are missing). Images have different voxel sizes: images of subjects pt001-pt009 have $0.65 \times 0.56 \times 4$ mm voxel size for CT images and $1.25 \times 1.25 \times 4$ mm for MRI images, while subjects pt101-pt109 have $0.45 \times 0.45 \times 3$ mm voxel size for CT images and $0.86 \times 0.86 \times 3$ mm for MRI images. Sample images and corresponding joint intensity distribution are shown in Figure 6.5.

Figure 6.5: *Sample images for CT to MRI-T1 registration and corresponding joint intensity distribution of registered images (darker color corresponds to higher probability).*

Table 6.1: *Registration errors for registering CT images to MRI-T1 images, for 10 anatomical points (VOI 1-10) and all 16 available patients.*

	VOI 1	VOI 2	VOI 3	VOI 4	VOI 5	VOI 6	VOI 7	VOI 8	VOI 9	VOI 10
pt001	1.318411	1.414050	2.067559	2.050089	1.539542	1.768205	1.951642	2.251120	1.451773	1.731921
pt002	———	1.361990	0.922660	0.659307	0.900857	0.512490	0.483710	1.693390	2.055826	1.921117
pt003	———	1.353354	1.572636	2.027756	1.402924	2.033294	1.336319	2.040582	1.560603	1.406557
pt004	———	———	———	———	———	1.961030	2.362611	1.749931	3.096772	3.112965
pt005	2.289893	2.194829	1.687327	1.396784	1.912644	1.477669	1.767082	2.069966	2.456541	2.473416
pt006	1.015653	1.037236	1.778054	2.088307	1.218297	1.646424	1.425881	1.834135	1.144019	1.278415
pt007	———	1.010118	1.003693	1.099953	1.000694	1.076686	1.021505	1.144803	1.015022	1.074432
pt101	2.214423	2.149883	2.050327	2.163085	2.118768	2.161599	2.022113	1.780373	2.098204	2.085105
pt102	1.362441	1.376753	2.389910	1.554298	1.580498	1.010325	2.415886	2.102256	1.308687	1.701420
pt103	2.597908	2.425481	2.369233	2.718541	2.480992	2.479410	2.021983	1.202205	2.244241	2.285358
pt104	1.559718	1.428977	1.343269	1.500759	1.491924	1.307704	1.097213	1.045764	1.617440	1.564693
pt105	2.355182	2.300441	2.360760	2.145938	2.283106	2.047241	2.451419	2.111519	2.220994	2.397404
pt106	1.822191	1.900953	2.066917	2.207116	1.915152	2.183549	1.879381	2.401565	2.055893	1.901160
pt107	2.504540	2.179510	2.085615	1.809047	1.802098	2.111907	3.040904	2.286672	1.897290	2.734406
pt108	1.422902	1.421461	1.451562	1.393379	1.422370	1.372511	1.469269	1.411851	1.406748	1.438731
pt109	1.621355	1.494555	1.463794	1.536585	1.541295	1.373264	1.296285	0.841338	1.660091	1.655860

mean = 1.76 mm, median = 1.76 mm, maximum = 3.11 mm

The results of registering CT images to MRI-T1 images are tabulated in Table 6.1. All the registrations were successful, resulting in overall mean error 1.76 mm, median error 1.76 mm and maximal error 3.11 mm. Results for subjects pt101-pt109, where images have smaller voxel sizes, were in general not better than the results for subjects pt001-pt107, as one may expect. This indicates that the source of registration errors is not in the discrete nature of data, but in the images themselves. Note that imaged anatomies are not absolutely rigid and furthermore, MRI images may be deformed due to the magnetic field inhomogeneity, caused by presence of the subject.

PET to MRI registration

Registration of PET images to MRI-T1 images was performed using 7 available CT-MRI image pairs (patients pt001-pt009, excluding subjects pt003 and pt004 where PET images were missing). The voxel size for PET images is $2.6 \times 2.6 \times 8$ mm, while the voxel size for corresponding MRI-T1 images is $1.25 \times 1.25 \times 4$ mm. Sample images and corresponding joint intensity distribution are shown in Figure 6.6.

The results of registering PET images to MRI-T1 images are summarized in Table 6.2. One can observe that registration error for subject pt006 is high for all 10 anatomical points, which indicates that this registration was not successful. Including this subject, the overall mean error of 3.58 mm, median error of 3.00 mm and maximal error of 11.22 mm were obtained. Excluding the wrong registration, which could be detected by visual inspection of registered images, the mean error is 2.55 mm, median error is 2.47 and maximal error is 5.47 mm.

Errors for registration of PET images are in general higher than errors for registration of CT images. There are two reasons: first, the resolution of PET images is much lower than resolution of CT images, and second, PET is a functional imaging technique such that PET images comprise relatively low amount of anatomical information, required for matching with MRI images.

Comparison with other systems

We have compared the results of our method with results obtained by other research groups, see [1]. To the date of our analysis, there were 43 participating groups, which contributed 113 sets of results. For the comparison we used only result sets with both CT-MRI T1 and PET-MRI T1 registration results, each of them obtained from registering at least 7 patients. Among these result sets we selected only the best one from each research group, and so we finally got 24 comparable result sets. We have compared them according to obtained mean registration errors. Comparison results are shown in Figure 6.7. Dashed lines and

Figure 6.6: *Sample images for PET to MRI-T1 registration and corresponding joint intensity distribution of registered images (darker color corresponds to higher probability).*

Table 6.2: *Registration errors for registering PET images to MRI-T1 images, for 10 anatomical points (VOI 1-10) and all 7 available patients.*

	VOI 1	VOI 2	VOI 3	VOI 4	VOI 5	VOI 6	VOI 7	VOI 8	VOI 9	VOI 10
pt001	0.974484	0.750237	1.740664	2.169446	0.629607	1.211050	0.799778	1.556204	1.089950	1.224679
pt002	———	1.699095	2.596562	2.353774	1.922822	1.812047	2.206658	1.043201	1.239837	1.479047
pt005	1.877124	1.815363	2.998484	2.266538	1.613928	1.537461	2.743053	2.265476	2.710502	2.333102
pt006	9.732683	9.593348	8.936985	6.138780	8.863224	6.246626	9.949534	9.149637	9.497526	11.219258
pt007	———	3.416825	5.472465	4.826450	3.864866	3.477531	4.631028	1.619964	2.593110	3.184805
pt008	———	3.537144	3.278741	3.701926	3.469292	3.755874	3.007465	3.104460	3.787213	3.452140
pt009	———	———	———	———	———	3.509255	4.634637	3.127904	3.072052	3.537593

mean = 3.58 mm, median = 3.00 mm, maximum = 11.22 mm

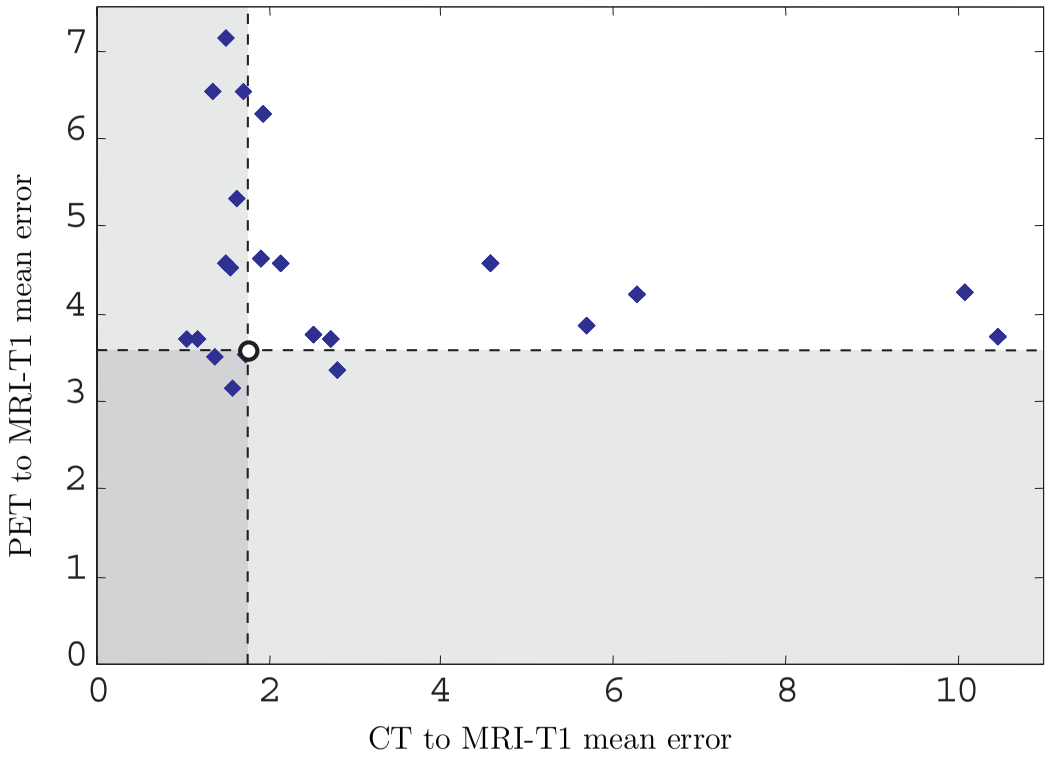


Figure 6.7: *Graph showing mean errors of different rigid registration systems, for CT to MRI-T1 and PET to MRI-T1 registration. Dashed lines correspond to mean values obtained by our rigid registration system, while other marks are used to mark best results of other research groups. The darker shading correspond to a region with better results than ours for both types of registration.*

Table 6.3: *Registration errors obtained by technique proposed by Collignon, Maes, Delaere, Vandermulen, Stuens and Marchal (CO), technique proposed by Studholme, Hill and Hawkes (ST) compared with results obtained by our registration system (RO).*

	CT to MRI-T1			PET to MRI-T1		
	mean	median	max	mean	median	max
CO	1.90	1.53	6.69	4.63	3.64	12.73
ST	1.36	1.17	2.78	3.50	3.25	9.32
RO	1.59	1.55	3.11	3.58	3.00	11.22

All results are in mm.

a circle correspond to mean values obtained by our rigid registration system, while other marks denote best results of other research groups. The darker shading corresponds to the region with better results for both types of registration while lighter shading corresponds to regions with better results in only one type of registration. Considering both types of registration, CT-MRI T1 and PET-T1, our system is one of the top six among 24 compared systems.

In [88] the Vanderbilt group, which governs the RIRE project, compared eleven different methods for rigid registration. Two of them use mutual information based similarity measures: a technique proposed by Collignon, Maes, Delaere, Vandermulen, Stuens and Marchal [18], and a technique proposed by Studholme, Hill and Hawkes [77]. In Table 6.3 we compare their results, (CO) and (ST), with results obtained with our system (RO). In the case of CT to MRI-T1 registration our method did not perform as good as the ST method, which performed best according to mean, median as well as maximal error, but better than the CO method. Similar results were obtained in the case of PET to MRI-T1 registration. Here our method and the ST method performed comparably well, ST method performed slightly better according to the mean error, while our method performed better according to the median error. The CO method did not perform as good as the other two methods. Here we have to note that in the case of PET to MRI-T1 registration all three methods misregistered one of the image pairs, which reflects in large maximal errors. The comparison of results obtained by CO and ST also show that the quality of registration highly depends on the system implementation.

The obtained results indicate that point similarity measures are comparable to mutual information measures in the case of rigid registration. Here, we have to point out that point similarity measures were designed for non-rigid registration tasks. Furthermore, the results show that our registration system is suitable for solving various multi-modality, not only non-rigid, but also rigid registration tasks.

6.3.2 Non-rigid registration

We present some experiments, which we have performed using our implementation of non-rigid registration. The purpose of the experiments was to aid to medical research. The experiments include mono-modality and multi-modality registration of 2D and 3D medical images.

Combining anatomical and functional information

By combining functional and anatomical images function can be related to structure. Intrinsically better resolution anatomical images, e.g. CT or MRI images, improve interpretation of the functional images, e.g. PET or fMRI images. However, the observed activity can be precisely mapped to the anatomical structures only if images are registered. In Figure 6.8 we present combining of anatomical information of thoracic CT images and functional information of corresponding PET images. Two types of PET images are available: emission image PET-em, which provides functional information, and transmission image PET-tr, which represents tissue attenuation and thus provides rough anatomical information. It is difficult to directly register PET-em image to CT image, while registration of PET-tr and CT images is easier. Because PET-tr and PET-em images are usually in register, PET-tr images are usually used for registration and the obtained transformation is used for combining PET-em images with CT images.

Brain activation experiments

Brain activation studies require averaging of functional fMRI (or PET) images of many individuals. The averaging improves the signal to noise ratio and the sensitivity of the method. However, since there is considerable individual variation in brain anatomy, images must be registered to the reference brain before being averaged. Because functional images do not comprise sufficient anatomical information they cannot be accurately registered to the reference image. To overcome the problem, the registration can be performed using anatomical MRI images, which are already in register with the functional images, and the obtained transformation applied to the functional fMRI images. For the illustration see Figure 6.9.

Computational morphometry

Computational morphometry of MRI images has become a standard tool in the statistical analysis of differences in brain structure between two groups of subjects. It searches for local differences, such as changes in gray matter density or spatial positioning. It requires spatial normalization of images of all the subjects in

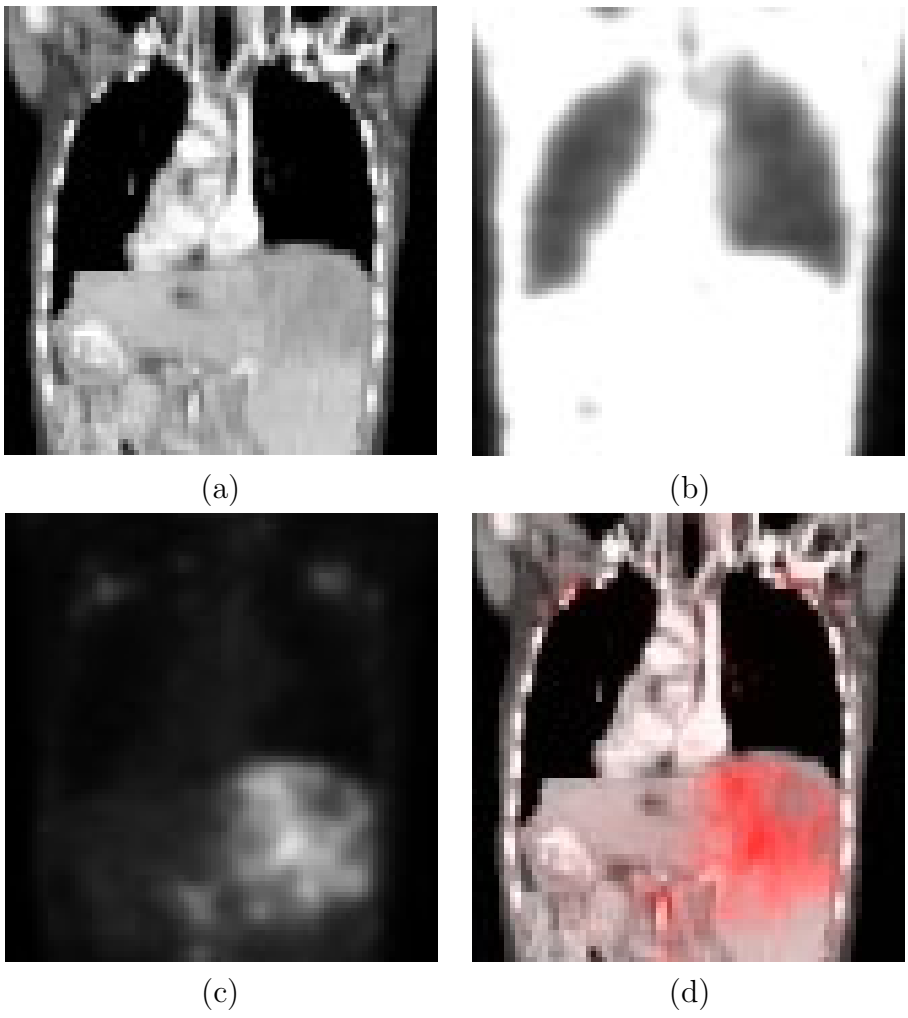


Figure 6.8: *Combining anatomical information of thoracic CT images and functional information of corresponding PET images. PET-tr (b) image is registerd to the CT image (a) and the obtained transformation is used for combining PET-em image (c) with the CT image (a). In the combined image (d) CT anatomical information is represented in grey while functional information of PET-em image is represented in red.*

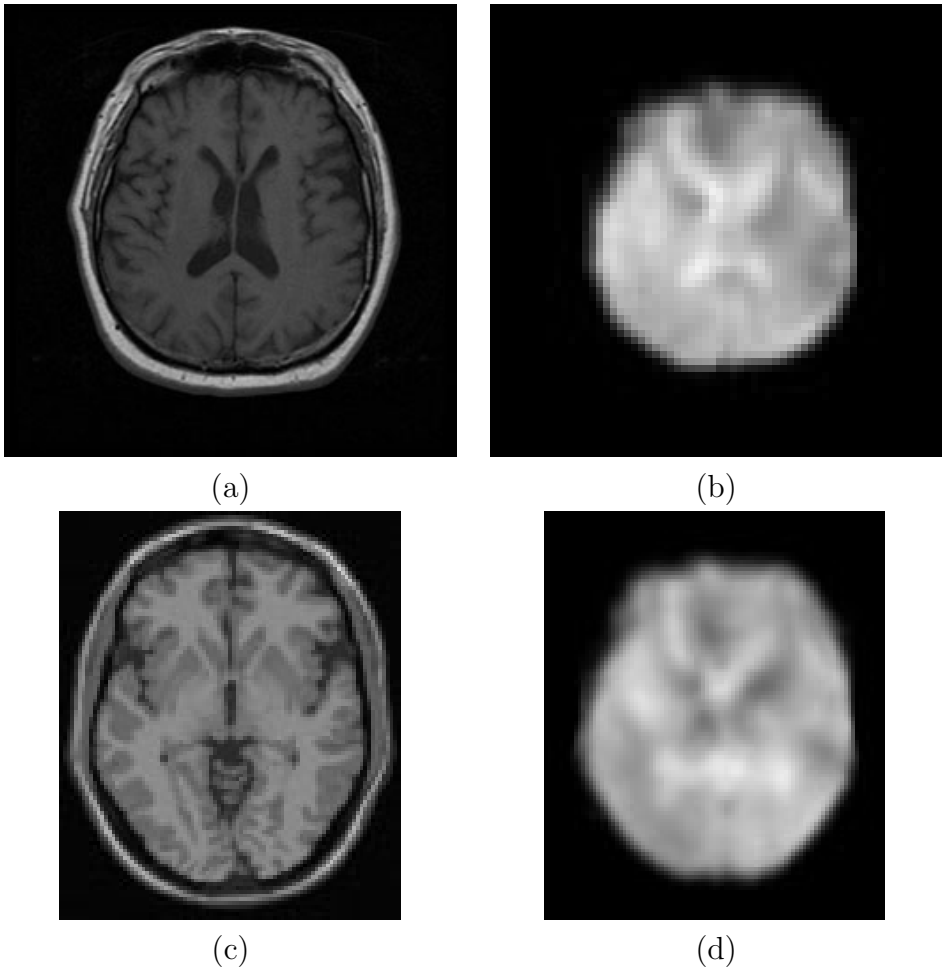


Figure 6.9: *Illustration of registration for brain activation studies. In order to put fMRI images (b) of many patients into correspondence, the corresponding MRI images (a) are registered to the reference image (c). The obtained transformations are applied to the fMRI images which thus get spatially aligned with the reference (d).*

the study into the same stereotactic space, which is performed by registering all the images to the same template. The normalized images are studied by analyzing differences in intensity values of the voxels (voxel-based morphometry VBM) [2] or by analyzing the deformation fields (deformation-based morphometry DBM and tensor based morphometry TBM) [15]. In Figure 6.10 we illustrate the registration for VBM, DBM and TBM with sample images and corresponding deformed image grid.

Recovering of breathing motion

MRI-T1 images of lungs were acquired successively during the subject was breathing. The aim of registration was to recover the non-rigid breathing motion by finding the correspondence between the successive images. The images have a voxel size of $2.73 \times 2.73 \times 10.00$ mm and the array size $128 \times 128 \times 1$ voxels. For the illustration see Figure 6.11, where two successive images and their overlap before and after the registration are shown. The corresponding transformation, which represents the assessed non-rigid breathing motion, is shown in Figure 6.12.

Finding an average anatomy

Registration can be used to find a mean shape of anatomical structures over several subjects. For the illustration we show a computation a mean shape of corpus callosum, see Figure 6.13. Images of several subjects were registered such that one of them served as a reference for registering all the others. The displacement maps obtained by the registration were averaged and used for deforming the reference image. The result of the deformation represents the average anatomy.

6.4 Conclusion

In this chapter we presented our system for medical image registration. It uses the approaches proposed in this thesis, i.e. point similarity measures, the symmetric registration approach and the combined spatial deformation model. The system was used to perform the experiments presented in this thesis, which prove the capabilities of the proposed methods. Furthermore we have used the system to perform several registrations intended for further medical research.

The system can be used for further research in the filed of medical image processing and also for the medical research. However it is not suitable for use in clinical medicine, because it is not yet suitably validated. However the system can be used to demonstrate capabilities of medical image registration in order to further promote the field of medical image processing.

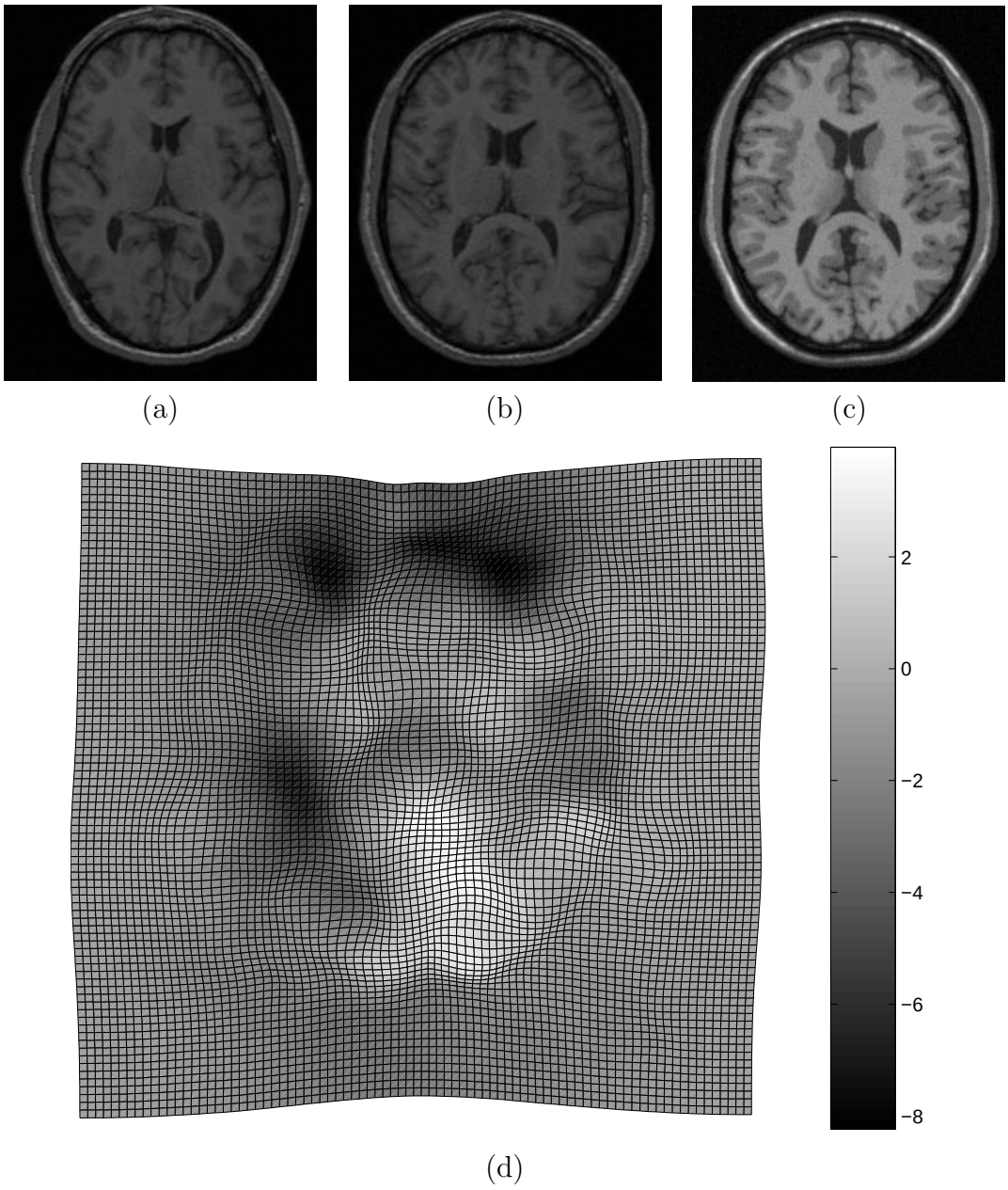


Figure 6.10: *Registration for computational morphometry. Each image from many subjects (a) is non-rigidly registered to the template (c), to obtain a normalized image (b). A slice of the deformed image grid, which is used for DBM and VBM, is shown in (d). Here, the grey level represents the deformation in the third spatial dimension.*

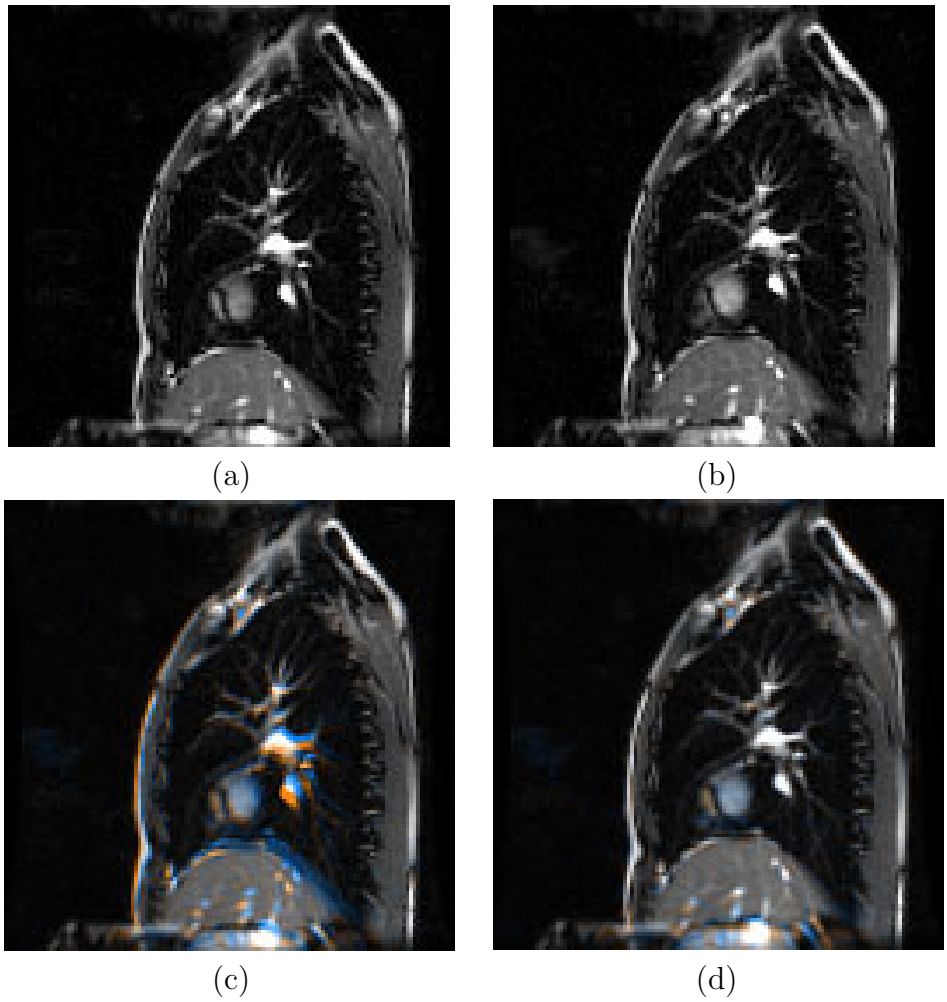


Figure 6.11: *Recovering of breathing motion. Each two successive images (a) and (b) were non-rigidly registered to assess the breathing motion. The image overlap before and after the registration is shown in (c) and (d) respectively. Here, the orange color represents higher intensity value in the first image, blue represents higher intensity values of the second image and grey represents equal intensity values in both of the images, indicating good image match.*

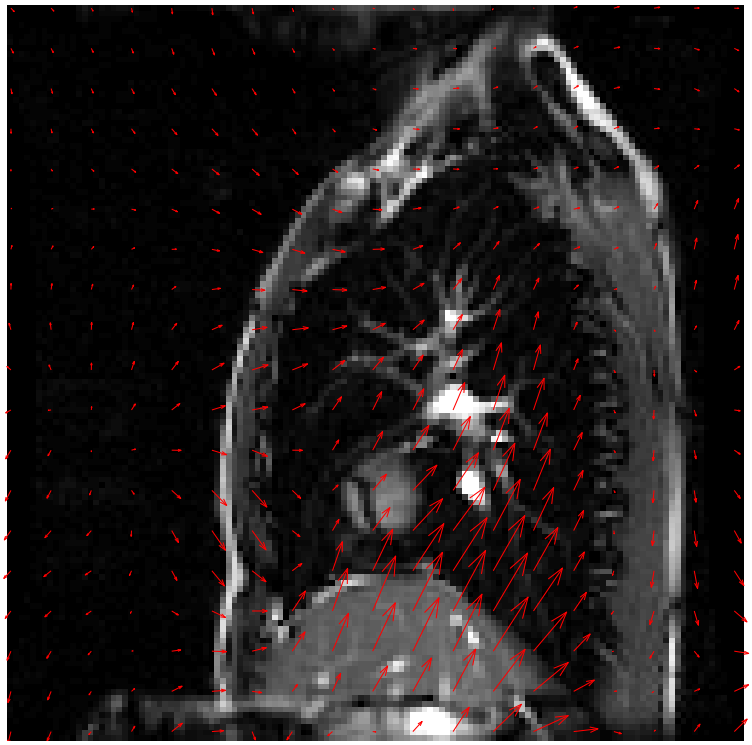


Figure 6.12: *The breathing motion assessed by non-rigid registration.*

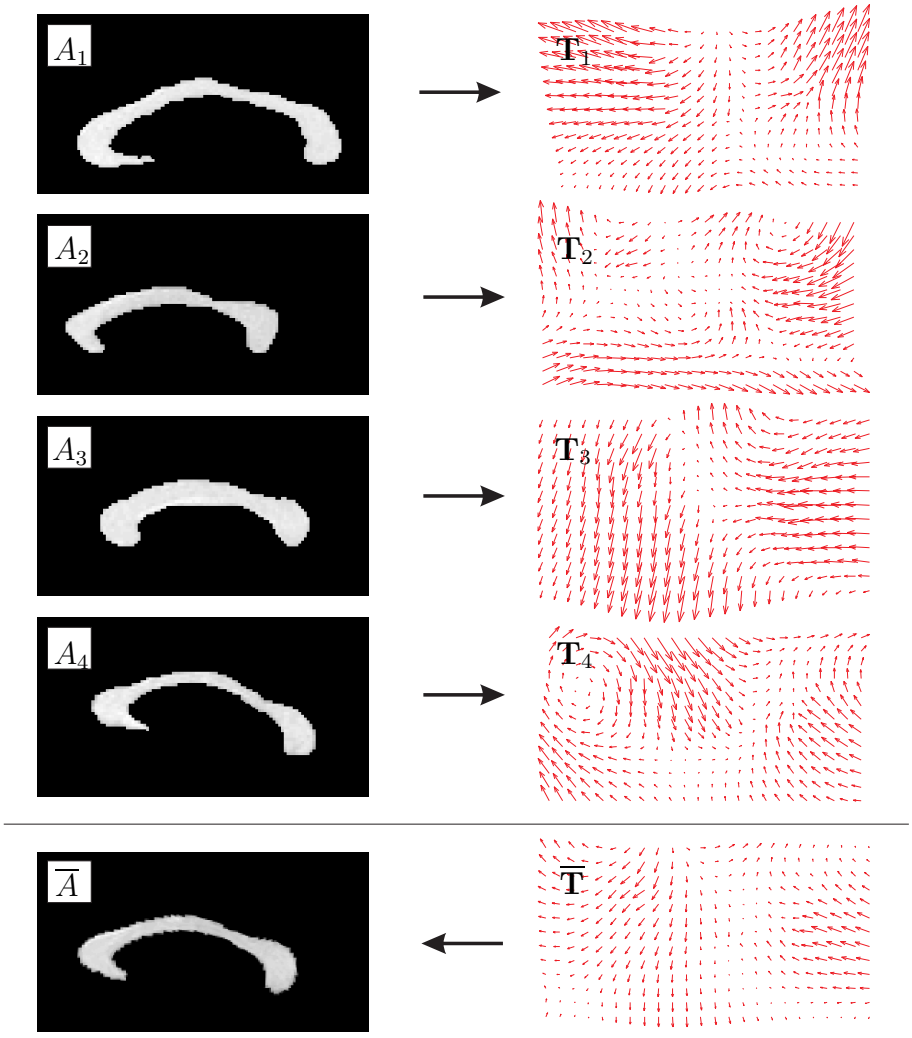


Figure 6.13: *Finding an average shape of corpus callosum. A reference image B was registered to images A_n of several subjects to obtain transformations $\mathbf{T}_n$ for each of them. The transformation $\overline{\mathbf{T}}$ was then computed by averaging transformations $\mathbf{T}_n$ and applied to the reference image B such that image of the average anatomy $\overline{A} = \overline{\mathbf{T}}B$ was obtained.*

7. Conclusion

In this thesis we focused on multi-modality non-rigid registration, which is one of the most difficult registration tasks. The main difficulty originates in detecting and localizing complex local multi-modality image discrepancies. Conventional multi-modality similarity measures, which are normally used to assess image discrepancies, are limited to large image regions and do not enable direct assessment of local image mismatches. Although these measures can be used to assess local image mismatches indirectly, by measuring global similarity at applied local transformations, they in practice limit the dimensionality of the registration by their high computational cost, which prevents detailed image registration. To solve this problem we devised a new group of similarity measures, which we call point similarity measures. They separate the process of measuring similarity into two steps: estimation of image intensity dependence and actual measurement of image similarity. This approach enables direct measurement of multi-modality similarity of arbitrary large image regions, including individual image points. It also leads to other advantages: the ability to avoid interpolation artifacts, possible improvement of position of the point similarity maxima and low computational cost. As such they perfectly suit the requirements of multi-modality non-rigid registration and enable its additional development. Because point similarity measures do not use any spatial information, the overall registration can be separated into two functionally independent registration stages: the first one that drives the registration in order to improve image similarity, and the second one that regularizes the registration according to the expected spatial deformation properties of imaged tissues. Both of the stages benefit from the point similarity measures. The first stage can follow the proposed principle of the symmetric image registration, which improves consistency and registration correctness. The benefit of the second stage is in the functional separation of both registration stages, which allows more precise modeling of spatial properties of the registration. We have also devised a new combined elastic-incremental spatial deformation model, which reduces the registration error with respect to popular elastic and incremental models. Finally, point similarity measures also enable the three step evaluation strategy, which improves evaluation of the overall registration system.

We have built an image registration system, which uses the proposed methods: point similarity measures, the symmetric image registration and the com-

bined elastic-incremental spatial deformation model. The system enables rigid and non-rigid registration of mono-modality and multi-modality images. It was used for the experiments presented in this thesis, which show the capabilities of the proposed methods. The system was also used to perform several different registration tasks for medical research. This proves the generality of the system, which makes it suitable for medical research as well as for promotion of the field of medical image registration. However, the system is not suitable for clinical use, because it has not been validated as required fot that purpose.

The system could be adapted to specific applications by using additional task specific knowledge, which would improve the registration results. Integration of prior knowledge to point similarity measures is straightforward. Point similarity functions could be improved by different kinds of prior knowledge of actual intensity dependence, e.g. by priorly marking certain intensity pairs as true or false or by providing a rough estimation of intensity dependence between the images. However, point similarity measures could also contribute to improvement of spatial deformation models, which still offer high potentials for future research. For example, segmentation based point similarity measure gives an estimate of segmentation in each registration iteration, which could be used to automatically build or improve a biomechanical spatial deformation model.

The methods proposed in this thesis deal with general problems of high dimensional and multi-modality registration. They improve the capabilities and performance of multi-modality non-rigid registration and make it, in this aspect, comparable to the mono-modality registration. In addition, the methods can be adapted to specific registration tasks by using additional knowledge and can be applied to specific clinically important applications.

References

[1] The retrospective registration evaluation project. *http://www.vuse.vanderbilt.edu/ image/registration*.

[2] J Ashburner and K.J. Friston. Voxel-based morphometry - the methods. *NeuroImage*, 11(6):805–821, 2000.

[3] J Ashburner, P Neelin, D.L. Collins, A.C. Evans, and K.J. Friston. Incorporating prior knowledge into image registration. *NeuroImage*, 6(4):344–352, 1997.

[4] R. Bajcsy and S. Kovačič. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, April 1989.

[5] M. Bro-Nielsen. *Medical Image Registration and Surgery Simulation*. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, 1996.

[6] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. *Springer Lecture Notes in Computer Science*, 1131:267–276, 1996.

[7] C. Broit. *Optimal registration of deformed images*. PhD thesis, University of Pennsylvania, 1981.

[8] L.G. Brown. A survey of image registration techniques. *ACM Comput. Surveys*, 24(4):325–376, December 1992.

[9] T. M. Buzug and J. Weese. Improving DSA images with an automatic algorithm based on template matching and an entropy measure. *Computer Assisted Radiology*, pages 145–150, 1996.

[10] T. M. Buzug and J. Weese. Voxel-based similarity measures for medical image registration in radiological diagnosis and image guided surgery. *Journal of Computing and Information Technology*, 6(2):165–179, 1998.

[11] T. M. Buzug, J. Weese, C. Fassnacht, and C. Lorenz. Elastic matching based on motion vector fields obtained with a histogram based similarity measure for DSA-image correction. In H.U. Lemke, M. W. Vannier, and K. Inamura,

editors, *Computer Assisted Radiology and Surgery*, pages 139–144, Berlin, Germany, 1997.

[12] P. Cachier and D. Rey. Symmetrization of the non-rigid registration problem using invertion-invariant energies : Application to multiple sclerosis. In S. L. Delp, A. M. DiGioia, and B. Jaramaz, editors, *Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, volume 1935 of *Lecture Notes in Computer Science*, pages 472–481. Springer, 2000.

[13] G.E. Christensen and H.J. Johnson. Consistent image registration. *IEEE Transactions on Medical Imaging*, 20(7):568–582, July 2001.

[14] G.E. Christensen, R.D. Rabbitt, and M.I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, 1996.

[15] M. K. Chung, K. J. Worsley, T. Paus, C. Cherif, D. L. Collins, J. N. Giedd, J. L. Rapoport, and A. C. Evans. A unified statistical approach to deformation-based morphometry. *NeuroImage*, 14(3):595–606, 2001.

[16] L. Clarke, R. Vethuizen, M. Camacho, J. Heine, M. Vaidyanathan, L. Hall, R. Thatcher, and M. Silbiger. MRI segmentation: Methods and applications. *Magnetic Resonance Imaging*, 13(3):343–368, 1995.

[17] C. A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans. Brainweb: Online interface to a 3D MRI simulated brain database. In *NeuroImage, Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain*, volume 5, May 1997.

[18] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In Y. Bizais, C. Barillot, and R. Di Paola, editors, *Information processing in medical imaging 1995*, pages 263–274. Kluwer Academic, 1995.

[19] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3D multi-modality medical image registration using feature space clustering. In N. Ayache, editor, *CVRMed*, volume 905 of *Lecture Notes in Computer Science*, pages 195–204. Springer, 1995.

[20] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, June 1998.

[21] O Cuisenaire, J.-P. Thiran, B. Macq, C. Michel, A. De Volder, and F. Marques. Automatic registration of 3D MR images with a computerized brain atlas. In *Medical Imaging 1996: Image Processing*, volume 2710 of *Proc. SPIE*, pages 438–448. SPIE Press, Bellingham, WA, 1996.

[22] M Ferrant, A. Nabavi, B. Macq, F.A. Jolesz, R. Kikinis, and Warfield S.K. Registration of 3-D intraoperative mr images of the brain using a finite-element biomechanical model. *IEEE Transactions on Medical Images*, 20(12):1384–1397, 2001.

[23] Matthieu Ferrant, Simon K. Warfield, Arya Nabavi, Ferenc A. Jolesz, and Ron Kikinis. Registration of 3d intraoperative mr images of the brain using a finite element biomechanical model. In S. L. Delp, A. M. DiGioia, and B. Jaramaz, editors, *Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, volume 1935 of *Lecture Notes in Computer Science*, pages 19–28. Springer, 2000.

[24] T. Gaens, F. Maes, D. Vandermeulen, and Suetens P. Non-rigid multimodal image registration using mutual information. In W.M. Wells, A. Colchester, and S. Delp, editors, *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 1099–1106, MIT, Cambridge, MA, USA, October 1998. Springer-Verlag.

[25] MI Miller GE Christensen, RD Rabbit. A deformable neuroanatomy textbook based on viscous fluid mechanics. In Prince and Runolfsson, editors, *Proceedings of the 1993 Conference on Information Sciences and Systems*, pages 211–216. Johns Hopkins University, 1993.

[26] J. C. Gee, D. R. Haynor, M. Reivich, and R. Bajcsy. Finite element approach to warping of brain images. In M. H. Loew, editor, *Medical Imaging 1994: Image Processing*, volume 2167 of *Proc. SPIE*, pages 327–337. SPIE Press, Bellingham, WA, 1994.

[27] J. C. Gee, L. Le Briquer, C. Barillot, D. R. Haynor, and R. Bajcsy. Bayesian approach to the brain image matching problem. In M. H. Loew, editor, *Medical Imaging 1995: Image Processing*, volume 2434 of *Proc. SPIE*, pages 145–156. SPIE Press, Bellingham, WA, 1995.

[28] James C. Gee. On matching brain volumes. *Pattern Recognition*, 32(1):99–111, 1999.

[29] J.C. Gee, M. Reivich, and R. Bajcsy. Elastically deforming a three-dimensional atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17:225–236, 1993.

[30] G. Gerig, M. Jomier, and M. Chakos. Non-rigid multimodal image registration using mutual information. In W.J. Niessen and M.A. Viergever, editors, *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'2001*, volume 2208 of *Lecture Notes in Computer Science*, pages 616–523. Springer, October 2001.

[31] C. Gramkow and M. Bro-Nielsen. Comparison of three filters in the solution of the navier-stokes equation in registration. In *Proceedings of the Scandinavian Conference on Image Analysis – SCIA'97*, pages 795–802, 1997.

[32] A. Guimond, A. Roche, N. Ayache, and J. Meunier. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE Transactions on Medical Imaging*, 20(1):58–69, January 2001.

[33] A. Hagemann, K. Rohr, H.S. Stiehl, U. Spetzger, and J.M. Gilsbach. Biomechanical modeling of the human head for physically based, nonrigid image registration. *IEEE Transactions on Medical Imaging - Special Issue on Model-Based Analysis of Medical Images*, 18(10):875–884, 1999.

[34] N. Hata, T. Dohi, S.K. Warfield, W.M. Wells, R. Kikinis, and F.A. Jolesz. Multimodality deformable registration of pre- and intraoperative images for MRI-guided brain surgery. In W.M. Wells, A. Colchester, and S. Delp, editors, *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 1067–1074, MIT, Cambridge, MA, USA, October 1998. Springer-Verlag.

[35] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D.L. Collins, A. Evans, G. Malandain, N. Ayache, G.E. Christensen, and H.J. Johnson. Retrospective evaluation of inter-subject brain registration. *IEEE Transactions on Medical Imaging*, 22(9):1120–1130, September 2003.

[36] P. Hellier, C. Barillot, E. Memin, and P. Perez. Hierarchical estimation of a dense deformation field for 3-d robust registration. *IEEE Transactions on Medical Imaging*, 20(5):388–402, May 2001.

[37] R. M. Henkelman. Measurement of signal intensities in the presence of noise in MR images. *Medical Physics*, 12(2):232–233, 1985.

[38] Gerardo Hermosillo and Olivier Faugeras. Dense image matching with global and local statistical criteria: a variational approach. In *Proceedings of CVPR 2001, Kauai, Hawaii, December 9-14, 2001*, volume 1, pages 73–78, 2001.

[39] D. L. G. Hill and D. J. Hawkes. Across-modality registration using intensity-based cost functions. In I. Bankman, editor, *Handbook of Medical Image Processing*, pages 537–553. Academic Press, 1999.

[40] M. Holden, D. L. G. Hill, E. R. E. Dent, J. M Jarosz, T. C. S. Cox, T. Rohlfing, Goodey J., and D. J. Hawkes. Voxel similarity measures for 3-D serial MR brain image registration. *IEEE Transactions on Medical Imaging*, 19(2):94–102, February 2000.

[41] P. J. Kostelec, J. B. Weaver, and D. M. Healy Jr. Multiresolution elastic image registration. *Medical Physics*, 25(9):1593–1604, September 1998.

[42] R.K.-S. Kwan, A.C. Evans, and G. B. Pike. An extensible MRI simulator for post-processing evaluation. In *Visualization in Biomedical Computing (VBC'96)*, volume 1131 of *Lecture Notes in Computer Science*, pages 135–140. Springer-Verlag, May 1996.

[43] H. Lester and S. R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32(1):129–149, 1999.

[44] M.E. Leventon and W.E.L. Grimson. Multi-modal volume registration using joint intensity distributions. In W.M. Wells, A. Colchester, and S. Delp, editors, *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 1057–1066, MIT, Cambridge, MA, USA, October 1998. Springer-Verlag.

[45] B. Likar and F. Pernuš. Automatic extraction of corresponding points for the registration of medical images. *Medical Physics*, 26:1678–1686, 1999.

[46] B. Likar and F. Pernuš. Registration of serial transverse sections of muscle fibers. *Cytometry*, 37(2):93–106, 1999.

[47] B. Likar and F. Pernuš. A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*, 19:33–44, 2001.

[48] B. Likar, M.A. Viergever, and F. Pernuš. Retrospective correction of MR intensity inhomogeneity by information minimization. *IEEE Transactions on Medical Imaging*, 20(12):1398–1410, 2001.

[49] R. Lucht, M.V. Knopp, and G. Brix. Elastic matching of dynamic MR mammographic images. *Magnetic Resonance in Medicine*, 43:9–16, 2000.

[50] F. Maes, A. Collignon, D. Vandermeulen, G. MArchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, April 1997.

[51] F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.

[52] J. B. A. Maintz, H. W. Meijering, and M. A Viergever. General multimodal elastic registration based on mutual information. In K.M. Hanson, editor, *Medical Imaging 1998: Image Processing*, volume 3338 of *Proc. SPIE*, pages 144–154. SPIE Press, Bellingham, WA, 1998.

[53] J.B.A. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[54] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 1993.

[55] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. Nonrigid multimodality image registration. In M. Sonka and K.M. Hanson, editors, *Medical Imaging 2001: Image Processing*, volume 4322 of *Proc. SPIE*, pages 1609–1620. SPIE Press, Bellingham, WA, 2001.

[56] M.I. Miller, G. E. Christensen, Y. Amit, and U. Grenander. Mathematical textbook of deformable neuroanatomies. In *Proceedings of the National Academy of Sciences*, volume 90, pages 11944–11948, 1993.

[57] M. Ono, S. Kubic, and C. Abernathey. *Atlas of the Cerebral Sulci*. Georg Thieme Verlag, Stuttgart, 1990.

[58] E. Parzen. On the estimation of probability density function. *Annual Mathematical Statistics*, 33:1065–1076, 1962.

[59] X. Pennec, N. Ayache, and J.P. Thirion. Landmark-based registration using features identified through differential geometry. In I. Bankman, editor, *Handbook of Medical Imaging*, pages 499–513. Academic Press, September 2000.

[60] X. Pennec, P. Cachier, and N. Ayache. Understanding the demon's algorithm. In C.J. Taylor and A. Colchester, editors, *Proceedings of the 2nd International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*, volume 1679 of *Lecture Notes in Computer Science*, pages 597–605. Springer, 1999.

[61] D.L. Pham, C. Xu, and J.L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315–337, 2000.

[62] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19(8):809–814, August 2000.

[63] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Interpolation artefacts in mutual information based image registration. *Computer Vision and Image Understanding*, 77(2):211–232, 2000.

[64] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, August 2003.

[65] A. Roche, G. Malandain, and N. Ayache. Unifying Maximum Likelihood Approaches in Medical Image Registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging*, 11(1):71–80, 2000.

[66] A. Roche, G. Malandain, N. Ayache, and S. Prima. Towards a better comprehension of similarity measures used in medical image registration. In C.J. Taylor and A. Colchester, editors, *Proceedings of the 2nd International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*, volume 1679 of *Lecture Notes in Computer Science*, pages 555–566. Springer, 1999.

[67] P. Rogelj and S. Kovačič. Local similarity measures for multimodal image matching. In S. Loncaric, editor, *Proceedings of the first International Workshop on Image and Signal Processing and Analysis – IWISPA 2000*, pages 81–86. University Computing Center, University of Zagreb, 2000.

[68] P. Rogelj, S. Kovačič, and J.C. Gee. Validation of a nonrigid registration algorithm for multimodal data. In M. Sonka and J.M. Fitzpatrick, editors, *Medical Imaging 2002: Image Processing*, volume 4684 of *Proc. SPIE*, pages 299–307. SPIE Press, Bellingham, WA, 2002.

[69] P. Rogelj, S. Kovačič, and J.C. Gee. Point similarity measures for non-rigid registration of multi-modal data. *Computer vision and image understanding*, 92(1):112–140, October 2003.

[70] G. K. Rohde, A Aldroubi, and B. M. Dawant. Adaptive free-form deformation for inter-patient medical image registration. In M. Sonka and K.M. Hanson, editors, *Medical Imaging 2003: Image Processing*, volume 4322 of *Proc. SPIE*, pages 1578–1587. SPIE Press, Bellingham, WA, 2001.

[71] K. Rohr. Extraction of 3D anatomical point landmarks based on invariance principles. *Pattern Recognition*, 32(1):3–15, 1999.

[72] K. Rohr, H.S. Stiehl, R. Sprengel, W. Beil, T.M Buzug, J. Weese, and M.H. Kuhn. Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on Medical Imaging*, 20(6):526–534, 2001.

[73] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D.J. Hawkes. Non-rigid registration of breast MR images using mutual information. In W.M. Wells, A. Colchester, and S. Delp, editors, *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 1144–1152, MIT, Cambridge, MA, USA, October 1998. Springer-Verlag.

[74] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, August 1999.

[75] J. A. Schnabel, C. Tanner, A. D. Castellano-Smith, A. Degenhard, M. O. Leach, D. R. Hose, D. L. G. Hill, and D. J. Hawkes. Validation of nonrigid image registration using finite-element methods: application to breast MR images. *IEEE Transactions on Medical Imaging*, 22(3):238–247, February 2003.

[76] C. Studholme, Hill. D.L.G., and D.J. Hawkes. Using voxel similarity as a measure of medical image registration. In E. Hancock, editor, *Proceedings of the British machine vision conference – BMVC'94*, pages 235–244, 1994.

[77] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated 3D registration of truncated MR and CT images of the head. In *Proc. Brit. Mach. Vis. Conf.*, pages 27–37, September 1996.

[78] C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32(1):71–86, 1999.

[79] R. Szeliski and S. Lavallee. Matching 3-D anatomical surfaces with non-rigid deformations using octree-splines. *International Journal of Computer Vision*, 18(2):171–186, May 1996.

[80] J.-P. Thirion. Fast non-rigid matching of 3d medical images. In *Proceedings of the Conference on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, Baltimore, November 1995.

[81] J.-P. Thirion. Image matching as a diffusion process: An analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.

[82] P. Thompson and A. W. Toga. Warping strategies for intersubject registration. In I. Bankman, editor, *Handbook of Medical Image Processing*. Academic Press, 1999.

[83] P. Viola and W. Wells III. Alignment by maximization of mutual information. In *Proceedings of the 5th International Conference on Computer Vision*, pages 16–23, 1995.

[84] P. Viola and W. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 1997.

[85] P.A. Viola. *Alignment by maximization of mutual information, Ph.D. thesis.* MIT, MA Cambridge, 1995.

[86] Y Wang and L.H. Staib. Physical model-based non-rigid registration incorporating statistical shape information. *International Journal of Medical Image Analysis*, 4(1):7–20, 2000.

[87] S.K. Warfield, J. Rexilius, P.S. Huppi, T.E. Inder, E.G. Miller, W.M. Wells, G.P. Zientara, F.A. Jolesz, and R. Kikinis. A binary entropy measure to assess nonrigid registration algorithms. In W.J. Niessen and Viergever M.A., editors, *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001*, volume 2208 of *Lecture Notes in Computer Science*, pages 266–274. Springer, 2001.

[88] J. West, J.M. Fitzpatrick, M.Y. Wang, B.M. Dawant, C.R. Maurer, R.M. Kessler, R.J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. van den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P.F. Hemler, D.L.G. Hill, D. Hawkes, C. Studholme, J.B.A. Maintz, M.A. Viergever, G. Malandain, X. Pennec, M.E. Noz, .G.Q. Maguire, M. Pollack, C.A. Pelizzari, R.A. Robb, and R. Hanson, D. Woods. Comparison and evaluation of retrospective intermodality image registration techniques. *Journal of Computer Assisted Tomography*, 21(4):554–566, 1997.

[89] R.P. Woods. Validation of registration accuracy. In I. Bankman, editor, *Handbook of Medical Imaging Processing and Analysis*, pages 491–497. Academic Press, 1999.

[90] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.

Biography

Peter Rogelj was born on January 26, 1975 in Ljubljana. He finished a primary school in Vače and proceeded to the Secondary Electrotechnical School in Ljubljana. He finished a four year secondary school programme and enrolled University of Ljubljana, Faculty of Electrical Engineering in 1993. He attended a five year university programme Process Automation and Informatisation and graduated in 1998 with a diploma thesis on automatic reading of serial numbers of electricity meters. In the same year he started his postgraduate study at the same faculty and employed there as a junior researcher. His research focused to image processing and analysis. In 2001 he reached a master's degree with a thesis on multi-modality non-rigid registration of medical images. In addition to his work at Faculty of Electrical Engineering in Ljubljana, he also visited some other universities: in spring 1999 he studied one month at Czech Technical University in Prague, and in summer 2001 he was three months working at University of Pennsylvania in Philadelphia, USA.

Personal data

First and last name	Peter Rogelj
Date of birth	January 26, 1975
Born in	Ljubljana
Home address	Vače 96, 1252 Vače
Citizenship	Slovenian
MESS code of researcher	19230

Education

1993	Finished a Secondary Electrotechnical School in Ljubljana.
1993-1998	Study in a five year university programme at University of Ljubljana, Faculty of Electrical Engineering.
julij 1998	Reached a master's degree with a thesis on automatic reading of serial numbers of electricity meters.
april 1999	Study at Czech Technical University in Prague, in Center for Machine Perception,
1998-2001	Postgraduate master study at Faculty for Electrical Engineering, University of Ljubljana.
april 2001	Reached master's degree at Faculty for Electrical Engineering, University of Ljubljana with a thesis on multi-modality non-rigid medical image registration.
2001-2003	Doctoral study at Faculty for Electrical Engineering, University of Ljubljana.
2001	Three month scientific work at University of Pennsylvania, General Robotics, Automation, Sensing and Perception Lab (GRASP), Philadelphia, USA.

Publications

Primary source of bibliographic records: shared data base COBIB.SI

ARTICLES AND OTHER COMPONENT PARTS

1.01 Original Scientific Article

[1] ROGELJ, Peter, KOVAČIČ, Stanislav, GEE, James C., Point similarity measures for non-rigid registration of multi-modal data. *Computer Vision and Image Understanding*, 2003, vol. 92, no. 1, pp. 112–140. [COBISS.SI-ID 3897684]

[2] HASEGAWA, Ichiro, UEMATSU, Hidemasa , GEE, James C., ROGELJ, Peter, SONG, Hee Kwon, NAKATSU, Masashi, TAKAHASHI, Masaya, GEFTER Warren B., HATABU, Hiroto, Voxelwise mapping of magnetic resonance ventilation-perfusion ratio in a porcine model by multimodality registration: Technical note, *Academic Radiology*, 2003, vol. 10, no. 10, pp. 1091–1096. [ISSN 1076-6332]

1.08 Published Scientific Conference Contribution

[3] ROGELJ, Peter, KOVAČIČ, Stanislav. Odčitavanje serijske številke električnih števcev. In: ZAJC, Baldomir (ed.). *Zbornik sedme Elektrotehniške in računalniške konference ERK '98, 24. - 26. september 1998, Portorož, Slovenija.* Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, 1998, vol. B, pp. 301-304, graf. prikazi. [COBISS.SI-ID 1113428]

[4] ROGELJ, Peter, DOBRIŠEK, Simon, MIHELIČ, France. Govorni uporabniški vmesnik v industrijskem okolju. In: ZAJC, Baldomir (ed.). *Zbornik osme Elektrotehniške in računalniške konference ERK '99, 23. - 25. september 1999, Portorož, Slovenija.* Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, 1999, vol. B, pp. 233-236, graf. prikazi. [COBISS.SI-ID 1592148]

[5] ROGELJ, Peter, KOVAČIČ, Stanislav. Local similarity measures for multimodal image matching. In: LONČARIĆ, Sven (ed.). IWISPA 2000 : *Proceedings of the first international workshop on image and signal processing and analysis in conjunction with 22nd international conference on information technology interfaces, Pula, Croatia, June 14-15, 2000.* Zagreb: University Computing Center, University of Zagreb, 2000, pp. 81-86, graf. prikazi. [COBISS.SI-ID 1874004]

[6] ROGELJ, Peter, KOVAČIČ, Stanislav. Similarity measures for multimodal elastic matching. In: ERŽEN, Ida (ed.), ŠTIBLAR-MARTINČIČ, Draga (ed.). *Life sciences 2000 : programme, abstracts, Gozd Martuljek, September 28th to October 1st, 2000.* Ljubljana: Slovenian Society for Stereology and Quantitative Image Analysis, 2000, pp. 35. [COBISS.SI-ID 1960276]

[7] ROGELJ, Peter, KOVAČIČ, Stanislav. Mere podobnosti pri poravnavi slik. In: ZAJC, Baldomir (ed.). *Zbornik devete Elektrotehniške in računalniške konference ERK 2000, 21. - 23. september 2000, Portorož, Slovenija.* Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, 2000, vol. B, pp. 237-240, graf. prikazi. [COBISS.SI-ID 2025300]

[8] ROGELJ, Peter, KOVAČIČ, Stanislav. Similarity measures for non-rigid registration. In: LIKAR, Boštjan (ed.). *Computer vision : proceedings.* Ljubljana: Slovenian Pattern Recognition Society, 2001, pp. 82-91, graf. prikazi. [COBISS.SI-ID 2149204]

[9] ROGELJ, Peter, KOVAČIČ, Stanislav. Similarity measures for non-rigid registration. In: SONKA, Milan (ed.), HANSON, Kenneth M. (ed.). *Medical imaging 2001, Image Processing : 19-22 February 2001, San Diego, USA,* (Progress in biomedical optics and imaging, vol. 2, no. 27), (Proceedings of SPIE, vol. 4322). Bellingham (USA): SPIE, cop. 2001, zv. 1, pp. 569-578. [COBISS.SI-ID 2339924]

[10] ROGELJ, Peter, KOVAČIČ, Stanislav, GEE, James C. Validation of a non-rigid registration algorithm for multi-modal data. In: SONKA, Milan (ed.), FITZPATRICK, J. Michael (ed.). *Medical imaging 2002, Image Processing : 24-28 February 2002, San Diego, USA,* (Progress in biomedical optics and imaging, vol. 3, no. 22), (Proceedings of SPIE, vol. 4684). Bellingham (USA): SPIE, cop. 2001, zv. 1, pp. 299-307. [COBISS.SI-ID 2822484]

[11] ROGELJ, Peter, KOVAČIČ, Stanislav. Rigid multi-modal medical image registration based on point similarity measures. In: ZAJC, Baldomir (ed.). *Zbornik enajste mednarodne Elektrotehniške in računalniške konference ERK 2002, 23.-25. september 2002, Portorož, Slovenija.* Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, [2002], vol. B, pp. 337-340. [COBISS.SI-ID 3206996]

[12] ROGELJ, Peter, KOVAČIČ, Stanislav. Rigid multi-modality registration of medical images using point similarity measures. In: DRBOHLAV, Ondřej (ed.). *Computer vision - CVWW'03 : proceedings of the 8th Computer Vision Winter Workshop, February 3-6, 2003, Valtice, Czech Republic.* Prague: Czech Pattern Recognition Society, 2003, pp. 159-163. [COBISS.SI-ID 3398996]

[13] ROGELJ, Peter, KOVAČIČ, Stanislav. Symmetric image registration. In: In: Milan Sonka (ed.), J. Michael Fitzpatrick (ed.). *Medical Imaging 2003: Image Processing, February 15-20, 2003, San Diego, California, USA.* Bellingham (USA): SPIE, cop. 2001, vol. 5032, pp. 334-343. [ISBN 0-8194-4833-8]

[14] ROGELJ, Peter, KOVAČIČ, Stanislav. Point similarity measure based on mutual information. In: GEE, James C. (ed.), MAINTZ, J. B. Antoine (ed.), VANNIER, Michael W. (ed.). *Biomedical image registration : revised papers*, (Lecture notes in computer science, 2717). Berlin; Heidelberg; New York: Springer-Verlag, cop. 2003, pp. 112-121. [COBISS.SI-ID 3912788]

[15] ROGELJ, Peter, KOVAČIČ, Stanislav. Netoga poravnava večmodalnih medicinskih slik s točkovnimi merami podobnosti. In: ZAJC, Baldomir (ed.). *Zbornik dvanajste mednarodne Elektrotehniške in računalniške konference ERK 2003, 25. - 26. september 2003, Ljubljana, Slovenija, (Zbornik ... Elektrotehniške in računalniške konference ERK ..., 1581-4572).* Ljubljana: IEEE Region 8, Slovenska sekcija IEEE, 2003, pp. 489-492. [COBISS.SI-ID 3857492]

MONOGRAPHS AND OTHER COMPLETED WORKS

2.09 Master's Thesis

[16] ROGELJ, Peter. *Elastična poravnava medicinskih slik različnih modalnosti : magistrsko delo.* Ljubljana: [P. Rogelj], 2001. X, 84 pp., ilustr. [COBISS.SI-ID 2222676]

2.11 Undergraduate Thesis

[17] ROGELJ, Peter. *Odčitavanje serijske številke električnih števcev : diplomska naloga.* Ljubljana: [P. Rogelj], 1998. 51 f., ilustr. [COBISS.SI-ID 1293652]

Zahvala / Acknowledgements

Iskreno se zahvaljujem mentorju prof. dr. Stanislavu Kovačiču, ki me je vodil skozi podiplomski študij, katerega rezultat je tudi ta doktorska disertacija. Zahvalil bi se tudi sodelavcem iz fakultete za elektrotehniko, ki so veliko prispevali k ustvarjanju dobrega delovnega vzdušja. Še posebna zahvala gre Janezu Peršu, Franciju Lahajnarju in Boštjanu Likarju za številne izmenjave izkušenj in idej.

I am grateful to my co-supervisor prof. dr. James C. Gee for inviting me to University of Pennsylvania, Philadelphia, and for fruitful discussions and scientific cooperation.

Zahvaljujem se tudi vsem, ki so mi na dosedanji življenski poti pomagali, verjeli vame in mi polepšali marsikateri trenutek. Posebna zahvala gre mami Milki in očetu Marjanu, ki sta me na moji poti podpirala in spodbujala.

Peter

Ljubljana, november 2003.

Izjava

Izjavljam, da sem doktorsko disertacijo izdelal samostojno, v sodelovanju z mentorjem prof. dr. Stanislavom Kovačičičem in somentorjem Jamesom C. Geejem. Pomoč drugih sodelavcev sem navedel v zahvali.

Ljubljana, november 2003 Peter Rogelj