

Introdução à Análise de Movimento usando Visão Computacional

Relatório Interno

*Raquel Ramos Pinho
João Manuel R. S. Tavares
Miguel F. P. Velhote Correia*



Universidade
do Porto

Faculdade de
Engenharia

FEUP



FEUP – Faculdade de Engenharia da Universidade do Porto
LOME – Lab. de Óptica e Mecânica Experimental do INEGI – Inst. de Engenharia e Gestão Industrial
INEB – Inst. de Engenharia Biomédica, Laboratório Sinal e Imagem
Rua Dr. Roberto Frias, s/n 4200-465 PORTO – PORTUGAL

Junho 2004

Agradecimentos

O trabalho apresentado neste relatório foi possível graças à colaboração e apoio de várias pessoas e Instituições cujo primeiro autor gostaria de destacar e agradecer. Nomeadamente, o apoio do Prof. João Tavares e do Prof. Miguel Correia pela sua disponibilidade e orientação; a cedência de instalações e as condições de trabalho proporcionadas pelo LOME – Laboratório de Óptica e Mecânica Experimental, e o bom ambiente e companheirismo proporcionados pelos colegas que integram a equipa do LOME; a Bolsa de Doutoramento concedida pela FCT – Fundação para a Ciência e Tecnologia no âmbito do projecto POSI sob a referência SFRH/BD/12834/2003; e especialmente a paciência e incentivo dos pais e irmão.

Resumo

Pretende-se com este trabalho fazer uma introdução ao que tem vindo a ser realizado no domínio do seguimento e análise de movimento recorrendo a visão computacional.

Assim no primeiro capítulo deste relatório faremos referência aos vários tipos de movimento e analisaremos as fases que compõem um sistema comum de captura e análise de movimento, descrevendo sucintamente alguns trabalhos realizados nesta área.

Seguidamente, no segundo capítulo, faremos uma apresentação mais detalhada da área do seguimento e análise de movimento humano de corpo inteiro; nomeadamente, no reconhecimento da pose e do reconhecimento do andar e de gestos.

Finalmente, no terceiro e último capítulo, daremos ênfase à análise de imagem médica e exemplificaremos, sumariamente, algumas das suas aplicações.

Abstract

With this work we intend to introduce what has been done in the domain of tracking and motion analysis by using computational vision.

Therefore in the first chapter of this report we will refer the various types of motion, and analyse the steps that compose a general system of movement capture and analysis, by succinctly describing some works done in this field.

Then, in the second chapter we will do a more detailed study about the area of human entire body tracking and motion analysis; namely, in pose recognition and in the recognition of gait and gestures.

Finally, in the third and last chapter, emphasis will be given to the medical images analysis and we will summarily exemplify some of its applications.

Índice

1. INTRODUÇÃO À ANÁLISE DE MOVIMENTO.....	1
1.1 CLASSIFICAÇÃO DO MOVIMENTO.....	2
1.2 CAPTURA DE MOVIMENTO.....	2
1.2.1 <i>Estrutura de um Sistema de Processamento de Imagem</i>	3
1.3 PRÉ-PROCESSAMENTO DE IMAGEM.....	3
1.4 SEGMENTAÇÃO	3
1.4.1 <i>Dados Temporais</i>	4
1.4.2 <i>Dados Espaciais</i>	5
1.5 REPRESENTAÇÃO	8
1.5.1 <i>Representação baseada no objecto</i>	8
1.5.2 <i>Representação baseada em imagem</i>	9
1.6 SEGUIMENTO	10
1.6.1 <i>Estabelecimento de correspondências entre características</i>	11
1.6.2 <i>Modelos Probabilísticos</i>	12
1.6.2.1 <i>Filtragem de Kalman e suas variantes</i>	12
1.6.2.2 <i>Filtros de Partículas</i>	13
2. ANÁLISE DE MOVIMENTO HUMANO.....	14
2.1 CAPTURA DE MOVIMENTO.....	14
2.1.1 <i>Estrutura de um Sistema de Processamento</i>	15
2.2 RESTRIÇÕES	15
2.3 SEGUIMENTO	16
2.3.1 <i>Inicialização</i>	16
2.3.1.1 <i>Câmaras</i>	16
2.3.1.2 <i>Modelos</i>	16
2.3.2 <i>Modelação</i>	18
2.3.3 <i>Seguimento Múltiplo</i>	22
2.4 RECONHECIMENTO.....	24
2.4.1 <i>Estimação da Pose</i>	25
2.4.2 <i>Reconhecimento do Andar e de Gestos</i>	26
3. ANÁLISE DE IMAGEM MÉDICA.....	27
3.1 SEGMENTAÇÃO EM IMAGEM MÉDICA	29
3.2 ANÁLISE E SEGUIMENTO EM IMAGEM MÉDICA	32
BIBLIOGRAFIA.....	34

1. Introdução à Análise de Movimento

A análise de movimento por visão computacional tem vindo a desenvolver-se ao longo dos últimos anos, nomeadamente nos domínios da análise de movimento do corpo humano e de estruturas em imagem médica, que constituem duas das áreas mais activas da visão computacional [Kambhamettu, 1994], [Baumberg, 1995], [Gavrila, 1999], [Tavares, 2000].

O interesse na análise de movimento não é surpreendente, uma vez que do ponto de vista técnico este domínio é rico e desafiante, pela necessidade de segmentar, seguir e analisar estruturas que muitas vezes se alteram em termos de topologia, que geralmente envolvem movimento do tipo não rígido e frequentemente sofrem oclusão parcial ou até mesmo total [Gavrila, 1999]. Os recentes avanços tecnológicos, também contribuíram para a evolução neste domínio, ao ser permitida a captura, transferência e processamento de imagens em tempo real em sistemas de *hardware* cada vez mais usuais. A análise do movimento do corpo humano tem vindo a ser motivada pela vantagem de melhorar a interacção homem/máquina em diversas aplicações, tais como: a análise do desempenho atlético (ver por exemplo [Pers, 2002]), em circuitos de vigilância [Cucchiara, 2000], animações em realidade virtual [Gleicher, 1999], diagnósticos e estudos clínicos [Mcinerney, 1996], [Lakany, 1999], [Tavares, 2000], [Correia, 2001], [Moeslund, 2001], [Aggarwal, 1999]. A maior parte do trabalho desenvolvido no domínio de análise de movimento humano é, normalmente, inserido na área de seguimento de sujeitos e estimativa da sua pose [Moeslund, 2001].

A análise de movimento por visão computacional, também pode ser utilizada na área biomédica para realizar, por exemplo, o estudo do movimento do coração e do pulmão [Mcinerney, 1996], ou analisar o fluxo sanguíneo no sentido de fazer a triagem de pacientes e determinar a extensão das lesões cardíacas existentes [Tavares, 2000].

Algumas das aplicações práticas que exemplificam a utilização de imagens variáveis ao longo do tempo serão a análise do tráfego automóvel, a previsão das condições atmosféricas através do movimento das nuvens, o estudo do movimento dos lábios para a sua leitura, ou a análise das deformações de materiais. Mas a análise de movimento usando visão computacional também pode ser utilizada noutros domínios, como por exemplo, para acelerar o estudo comportamental das abelhas, uma tarefa tida até então como árdua e demorada [Feldman, 2003].

A análise do movimento é fundamentada, por exemplo, por estudos psicológicos que indicam que o movimento é um mecanismo de agrupamento visual primário; enquanto que a capacidade de detectar objectos a partir de características bidimensionais como a cor, textura ou forma parece desenvolver-se posteriormente [Spelke, 1994].

1.1 Classificação do Movimento

Comumente, o movimento diz-se rígido se a distância entre quaisquer dois pontos do objecto é preservada, isto é, o objecto não estica nem dobra, sendo preservadas a curvatura média e a curvatura *Gaussiana* da superfície [Tavares, 2000].

A hipótese de movimento rígido falha em numerosas situações de análise de movimento, uma vez que muitos objectos da vida real não são rígidos; na verdade, todos os objectos são mais ou menos deformáveis. Por exemplo as árvores balançam, o papel dobra-se, a roupa amarrota-se, etc. [Kambhamettu, 1994].

O movimento não rígido pode ser classificado de diversas formas; ver por exemplo [Kambhamettu, 1994]. Neste trabalho, utilizamos a classificação utilizada em [J Wang, 2003], segundo a qual o movimento não rígido pode ser articulado, elástico ou fluido. O movimento articulado ocorre se as partes rígidas de um objecto se movem independentemente das restantes. De acordo com a caracterização dos tipos de movimento supracitada, a maior parte do movimento do corpo humano é não rígido e rígido por partes, sendo por isso geralmente classificado como articulado. O movimento elástico caracteriza-se pelo movimento não rígido com algum grau de continuidade e suavidade, como será o exemplo do movimento do coração, o acenar de um lenço, o dobrar de uma folha de metal, onde a forma do objecto deforma segundo determinadas restrições. O movimento fluido é um movimento não rígido que não satisfaz a restrição de continuidade, podendo envolver variações topológicas e deformações turbulentas [Tavares, 2000].

1.2 Captura de Movimento

Dada a variedade de aplicações da análise de movimento, são inúmeras as técnicas de captura que podem ser utilizadas. Assim, para cada aplicação podem ser escolhidas as técnicas que melhor se coadunam às restrições e necessidades específicas da área.

Mas de uma forma geral, pode-se considerar que os tipos de imagens utilizados são gerados pela combinação de uma fonte de iluminação com a reflexão e absorção de energia por parte dos elementos da cena que são visualizados [R Gonzalez, 2003]. Os termos “iluminação” e “cena” referidos são bem mais gerais do que uma fonte de luz visível que ilumina uma cena 3D comum. Por exemplo, a iluminação poderá ser oriunda de uma fonte de energia electromagnética tal como um radar, infravermelho ou raio-X; ou então de fontes menos comuns como *ultrasons* ou *PET's* (*Positron Emission Tomography*). De modo similar, os elementos da cena podem ser objectos familiares, ou também poderão ser moléculas, rochas ou estruturas do cérebro humano. Dependendo da fonte de luz, a energia de iluminação é reflectida ou transmitida pelos objectos.

Por exemplo, na monitorização de insectos podem ser incluídos dispositivos ópticos e optoelectrónicos, câmaras, identificação por rádio-frequência, rádio-telemetria, etc. Estas técnicas podem ser utilizadas para gravar e analisar o movimento em túneis de vento. Ainda neste domínio,

também se pode recorrer a imagens aéreas ou por satélite, por exemplo para deduzir as alterações na distribuição da população em algumas espécies migratórias de insectos [Reynolds, 2002].

1.2.1 Estrutura de um Sistema de Processamento de Imagem

Um sistema típico utilizado para processamento de imagens de carácter geral, começa pela captação por parte de um dispositivo físico (sensor) de energia radiada pelo objecto que se pretende visualizar. Os dados recolhidos pelo sensor são de seguida convertidos por um digitalizador num formato digital. O tratamento dos dados é feito por intermédio de um computador cujo *software* executa tarefas específicas, e a exibição dos dados é feita através de monitores [R Gonzalez, 2003].

Independentemente do contexto do processamento de imagem em análise de movimento, três aspectos comuns podem ser identificados. Quase todos os problemas começam por um problema de segmentar o objecto em causa do resto da imagem. De seguida, as áreas segmentadas podem ser transformadas numa outra representação no sentido de reduzir a quantidade de informação. E por fim, é definido como o objecto deve ser seguido entre quadros, isto é ao longo da sequência de imagens.

1.3 Pré-Processamento de Imagem

Uma vez adquiridas as imagens, poderá ser necessário utilizar algumas técnicas de pré-processamento no sentido de se facilitar a análise dos dados recolhidos. Para tal, poder-se-ão utilizar técnicas de realce, de restauração, de processamento de imagens a cores ou compressão das imagens [R Gonzalez, 2003].

As técnicas de realce de imagem têm por objectivo sobressair detalhes obscuros, ou salientar características de interesse da imagem. Um exemplo familiar de realce de imagem consiste em aumentar o contraste.

A restauração de imagem também visa a melhoria da aparência. As técnicas utilizadas são geralmente baseadas em modelos probabilísticos ou matemáticos de degradação da imagem.

A utilização de imagens a cores tem vindo a ser cada vez mais frequente, até porque a cor pode ser utilizada para a extracção de características de interesse. Assim, há que averiguar se os métodos de escalas de cinzento se mantêm consistentes com as propriedades dos espaços de cor, ou se os referidos métodos devem ser reformulados.

A compressão de imagens consiste na utilização de técnicas para a redução do espaço de armazenamento necessário para salvar as imagens ou da largura de banda para as transmitir.

1.4 Segmentação

Feito o pré-processamento das imagens captadas, para prosseguir com a análise de movimento é necessário determinar as regiões que correspondem aos objectos na cena em causa. Este processo designa-se por segmentação.

Na segmentação de imagens baseada em modelos, são exploradas informações prévias acerca dos objectos e/ou da cena através da imposição de restrições, como por exemplo sobre as fronteiras das regiões a segmentar.

Neste trabalho fazemos a distinção entre a segmentação obtida com base em informação temporal ou espacial.

1.4.1 Dados Temporais

Em grande parte dos casos, o uso de dados temporais pressupõe a hipótese de o fundo da cena e também a câmara permanecerem estáticos. Assim, através da diferença entre imagens pode-se determinar o movimento do objecto em estudo. A determinação das referidas diferenças, pode ser realizada através da subtracção entre imagens ou pelo cálculo do fluxo óptico [Moeslund, 2001].

A subtracção de imagens é amplamente utilizada e caracteriza-se por subtrair a intensidade ou o gradiente de cada *pixel* entre várias imagens. O resultado reflecte o movimento e o ruído entre imagens a menos que o objecto tenha a mesma intensidade/cor do fundo da cena. Caso a cena seja estática, pode ser adquirida uma imagem da mesma sem ruído e sem qualquer objecto e utiliza-la como referência no esquema de subtracção. Uma versão mais avançada consiste em actualizar a imagem do fundo da cena durante o processamento [Haritaoglu, 1998]. A subtracção apresenta bons resultados em situações controladas; contudo para imagens reais de ambientes exteriores, detectam-se várias limitações devido a possíveis variações de iluminação, da existência de sombras e de reflexos, da presença de contraste reduzido e/ou da possibilidade de ocorrer oclusão que necessariamente degradam os resultados obtidos [Correia, 1995].

Por outro lado, o fluxo óptico [Correia, 2001] baseia-se no movimento coerente de pontos ou características entre imagens e pode ser utilizado para fazer a segmentação em imagens. Por exemplo em [Yamamoto, 1991] os parâmetros do movimento de uma parte do corpo humano podem ser estimados a partir do fluxo óptico dos vários pontos que constituem essa parte do corpo, através da comparação com o movimento do modelo pré definido. Em [Bregler, 1997], assim como em [Bregler, 1998], cada *pixel* é representado pelo seu fluxo óptico, sendo agrupados em objectos binários de grande dimensão (designados por *blobs*) com movimentos coerentes e representados por uma combinação de *Gaussianas* multivariadas. Em [Gu, 1994] são segmentadas as orlas de intensidade em imagens consecutivas pelo seu comprimento e contraste também utilizando fluxo óptico.

Na segmentação baseada em dados temporais, e consequentemente nos exemplos mencionados, assume-se que o único objecto que se move em cena é o objecto em estudo, e portanto as alterações das intensidades dos *pixels* devem-se apenas à existência de movimento. Em muitos casos, os dados temporais são uma forte alternativa aos dados espaciais pois, geralmente, são mais fáceis de extrair.

1.4.2 Dados Espaciais

A segmentação com recurso a dados espaciais pode ser realizada por binarização ou segundo abordagens estatísticas. A segmentação por binarização consiste num processo simples baseado em hipóteses relativas à cena ou ao objecto. A segmentação por abordagens estatísticas é uma classe mais avançada, em que são relaxadas algumas hipóteses de aparência exploradas pelos métodos de subtracção.

Caso a cor/intensidade do objecto possa ser distinguida do resto da cena, poder-se-ão segmentar imagens por binarização como em [Darrell, 1994] e em [Iwai, 1999]. Outra versão consiste em utilizar marcas passivas ou activas no objecto em estudo que sejam facilmente segmentadas por binarização [Campbell, 1995], [Goncalves, 1998]. Uma abordagem semelhante recorre à utilização de câmaras de infra-vermelhos e marcas especiais detectadas por estas [Iwasawa, 1997].

As abordagens estatísticas utilizam as características individuais dos *pixels* ou de grupos de *pixels*, tais como cor e contornos, para extrair o objecto da imagem. Algumas destas abordagens (estatísticas) consideram uma sequência de imagens do fundo da cena e calculam a média e variância da intensidade/cor de cada *pixel* ao longo do tempo. Em cada imagem o *pixel* é comparado com as estatísticas da imagem de fundo e classificado como pertencente ou não ao mesmo. O problema da detecção de movimento pode ser implementado para que a diferença entre quadros/imagens seja modelada como uma mistura de duas distribuições *Gaussianas*. Este modelo não é adequado quando o fundo da cena tem orlas de intensidade junto aos objectos, mas é muito rápido [J Wang, 2003]. Por exemplo, em [Stauffer, 1999] é utilizada uma mistura de *Gaussianas* adaptativas para modelar o fundo, depois cada *pixel* é classificado como pertencente ou não a este a partir da distribuição *Gaussiana* que melhor o representa, segmentando-se assim de forma robusta a região de movimento mesmo em ambientes reais (fig. 1). Também, em [Cucchiara, 2000] é utilizada uma abordagem estatística em circuitos de vídeo-vigilância para segmentar rapidamente objectos em movimento em ambientes naturais não controlados. Em [Wren, 1997] analisa-se o movimento humano, sendo o sujeito modelado por um conjunto de *blobs* com cores e estatísticas espaciais individuais, e cada *pixel* da imagem é classificado como pertencente a um dos *blobs* de acordo com as suas propriedades espaciais e de cor.

Outra abordagem estatística consiste em utilizar contornos estáticos ou dinâmicos. Os contornos estáticos, referem-se ao uso pré-definido de estruturas estáticas que representam bordão contorno do objecto, tais como segmentos. Os contornos dinâmicos, ou activos, são utilizados de forma semelhante exceptuando o facto de se poderem ajustar aos objectos [Kass, 1988], [Cootes, 1992] e geralmente funcionam adequadamente quando a estrutura destes é desconhecida.

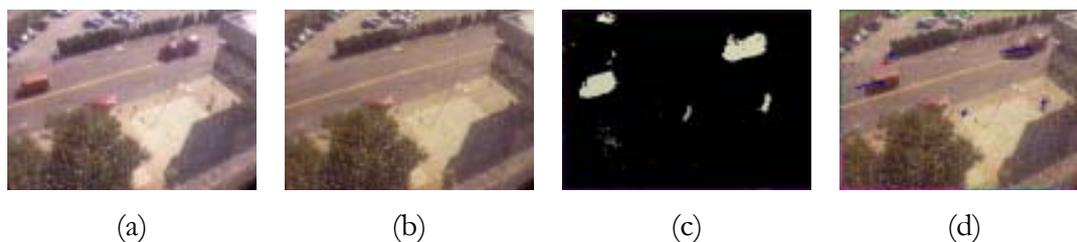


Fig. 1. Resultados da segmentação pela abordagem apresentada em [Stauffer, 1999] onde (a) é a imagem inicial; (b) é a imagem composta pelas médias das *Gaussianas* mais prováveis no modelo do fundo da cena; (c) segmentação da imagem dada realçando os *pixels* dos objectos a seguir; (d) a imagem inicial com a informação de seguimento sobreposta.

Os já referidos contornos activos, um caso particular da teoria multidimensional de modelos deformáveis [Mcinerney, 1996], têm vindo a receber atenção crescente ao longo dos últimos anos na área da visão computacional, pela sua versatilidade e flexibilidade para a modelação e representação de objectos [Jain, 1998]. Um modelo deformável é activo no sentido de ser capaz de se adaptar aos dados, conciliando as restrições geométricas da forma com as evidências locais da imagem. O trabalho desenvolvido na área de modelos deformáveis pode ser subdividido entre duas classes: os modelos de forma livre ou os modelos deformáveis paramétricos. Os primeiros, também designados por *snakes* ou contornos activos, representam quaisquer formas arbitrárias desde que seja satisfeita alguma restrição de regularidade (como continuidade ou suavidade). Por seu lado, nos modelos deformáveis paramétricos, a forma é caracterizada por uma fórmula paramétrica ou utilizando um protótipo e os seus modos de deformação. Estes modelos paramétricos são geralmente utilizados quando existe alguma informação prévia sobre a forma geométrica, e a mesma é representada por um conjunto preferencialmente reduzido de parâmetros [Jain, 1998].

Um exemplo da utilização *snakes* pode ser encontrado em [Hanek, 2004], onde é realizada a segmentação de imagens, baseada no algoritmo de contracção da densidade de curva (*Contracting Curve Density*). O algoritmo pode ser empregue para fazer o seguimento de estruturas ao longo de imagens. Para tal, utiliza-se uma função de aproximação entre os dados e a curva modelo, que recorre a modelos suavizados para determinar a densidade posterior, sendo a densidade de curva na imagem contraída no sentido de se obter uma curva simples (fig. 2). De acordo com os autores, o seguimento obtido por esta técnica supera todos os métodos de seguimento em precisão, robustez e tempo de execução.

Em [Zaritsky, 2003] o seguimento de objectos realiza-se com contornos activos 3D que incorporam informação acerca da velocidade local com bons resultados de seguimento, executando apenas uma iteração por quadro e portanto sem grande custo computacional.

Outro exemplo da utilização de modelos deformáveis pode ser encontrado em [Nascimento, 2003b], onde é proposto um novo algoritmo para a estimação da orla de objectos na presença de “*outliers*”. Essa estimação é feita através de um contorno deformável semelhante aos contornos activos, mas cuja função de potencial se adapta ao longo do processo, permitindo resultados mais robustos no caso de cenas com fundos complexos (fig. 3).



Fig. 2. Resultado da segmentação ao longo das iterações de acordo com a abordagem apresentada em [Hanek, 2004].

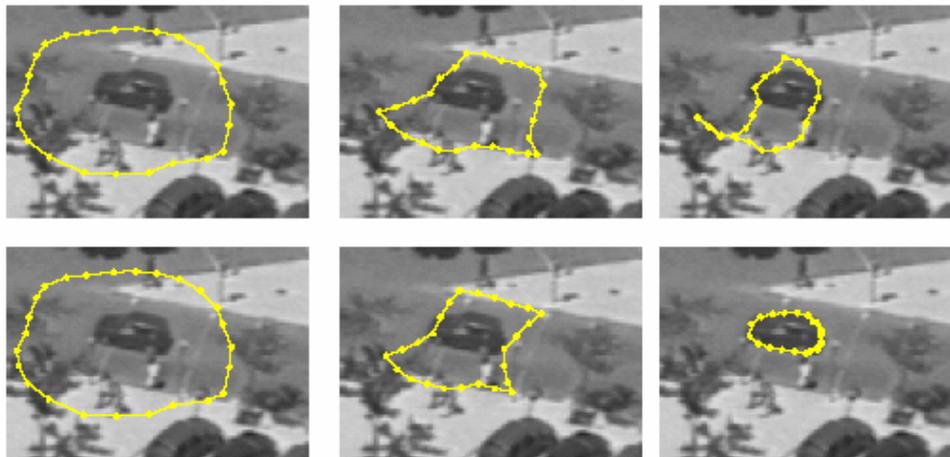


Fig. 3. Resultados obtidos pelos contornos activos usuais (na primeira linha) ou com função de potencial adaptativo (na segunda linha) conforme proposto em [Nascimento, 2003b].

Em [Bamford, 1998] é apresentada uma técnica de segmentação automática baseada na utilização de contornos activos duais, feita através da minimização da energia global e da aplicação do algoritmo de *Viterbi*. A sua abordagem é utilizada para segmentar núcleos estáticos de células, e parte do princípio que o contorno a aproximar é suave.

Outro exemplo poderá ser encontrado em [Xie, 2004], onde é apresentado um algoritmo de contornos activos guiados por restrições de regiões. A proposta feita é robusta ao ruído e a orlas de intensidade fracas, permitindo também a detecção de objectos múltiplos (fig. 4).

Também em [Wang, 2004] é proposta uma abordagem que permite fazer a segmentação de imagens mesmo na presença de objectos com fronteira não muito bem definida, complexa ou de forma irregular, e com interior heterogéneo. Esta abordagem integra informação relativa à região delimitada pelo contorno deformável, quer ao nível da homogeneidade, da textura ou da cor.

Uma das aplicações da utilização de contornos activos poderá ser a extracção de silhuetas. Por exemplo, em [Niyogi, 1994] as silhuetas são extraídas por um método robusto que utiliza contornos deformáveis

em tempo e espaço. Uma abordagem semelhante de extracção da silhueta é proposta em [Rigoll, 2000] que utiliza pseudo modelos de *Markov* não observáveis. Em [Garcés, 1999] é comparado e combinado o uso de contornos activos, com o fluxo óptico, e o seguimento 3D, no sentido de se obterem melhores resultados na extracção de silhuetas.

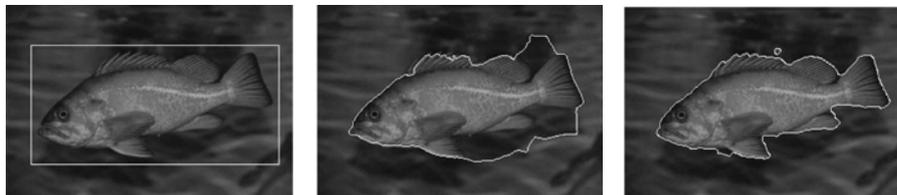


Fig. 4. *Snake* inicial e os resultados da *snake* geométrica *standard* e *snake* geométrica guiada por regiões proposta em [Xie, 2004].

A binarização no processamento de dados espaciais é dependente de várias hipóteses sobre a aparência do objecto e da cena em causa. Em aplicações com menos restrições, devido à sua adaptabilidade os métodos estatísticos são adequados; sendo que se obtém resultados mais fiáveis quando se utilizam métodos sobre regiões e não tanto individualmente sobre *pixels*, ainda que seja mais complexo modelar entidades de nível mais elevado. Os contornos activos podem ser eficientes quando têm uma boa estimativa inicial, mas apresentam algumas dificuldades. Por exemplo, as inconveniências decorrentes da segmentação de objectos articulados podem ser superadas ao utilizar um contorno activo para cada entidade dos referidos objectos. Contudo, de uma forma geral as abordagens estatísticas têm vindo a ser cada vez mais utilizadas devido à sua robustez quando comparadas com as abordagens de subtracção [Moeslund, 2001].

1.5 Representação

As entidades segmentadas podem ser descritas de forma compacta por alguma representação conveniente. Existem dois tipos principais de representação: as representações baseadas nos objectos, que têm por base a segmentação obtida por binarização, e as representações baseadas nas imagens que são determinadas directamente a partir das mesmas [Moeslund, 2001].

Nas secções seguintes faremos uma análise mais detalhada das referidas formas de representação.

1.5.1 Representação baseada no objecto

Este tipo de representação depende essencialmente dos dados obtidos pela segmentação da estrutura, e pode ser obtida considerando-se pontos, caixas, contornos ou áreas.

A representação por pontos é suficiente em sistemas usando marcadores passivos ou activos nos objectos a seguir. Os marcadores activos originam um elevado contraste entre imagens e providenciam uma representação robusta [Silaghi, 1998]. Caso sejam utilizadas duas ou mais câmaras, pode-se obter uma representação 3D da estrutura.

A representação por uma caixa é utilizada em muitos sistemas. A ideia consiste em representar o objecto a seguir por um conjunto de caixas que limitam *pixels*/zonas encontradas pela segmentação.

A representação por contornos é popular pela sua simplicidade. Por exemplo, no caso do movimento humano os contornos podem delimitar as silhuetas obtidas por segmentação a partir de binarização ou métodos de subtracção. A representação 2D é geralmente fácil de obter, mas também pode ser mais complexa como em [Baumberg, 1995].

Na representação por áreas, o objecto a seguir é representado por uma ou várias áreas com características similares, por exemplo coerentes com o fluxo e/ou cor.

1.5.2 Representação baseada em imagem

Os dados obtidos por segmentação podem ser representados directamente a partir dos *pixels* que compõem a imagem. Assim, as orlas de intensidade podem ser representadas por pontos, segmentos de recta ou por outras características [Christensen, 1997].

No sentido de compactar a representação dos dados/imagens, estes podem ser transformadas noutros espaços ou as suas representações expandidas por funções de base não cartesiana, como por exemplo de *Fourier* ou análise de componentes principais [Moeslund, 2001].

Por exemplo em [Cootes, 1995] é utilizada a análise de componentes principais, obtendo assim uma representação compacta que melhorou a eficiência e permitiu a generalização do conjunto de treino utilizado na metodologia. A análise de componentes principais é uma ferramenta estatística comum para reduzir a dimensão associada aos dados [Sidenbladh, 2000], e extrair características importantes destes em termos da quantidade de variância retida [Jain, 1998], desempenhando um papel importante na aprendizagem de representações de objectos. Contudo, geralmente estes métodos não permitem a ocorrência de oclusões; são necessários bons conjuntos de treino para que seja possível obter resultados adequados; e as parametrizações escolhidas poderão incluir estados cuja interpretação física não seja plausível [Gavrila, 1999]. Também, em [Franke, 1998] é utilizada a análise de componentes principais na representação de peões por grelhas; assim através da representação compacta do conjunto de treino a partir de vectores próprios e das grelhas, a oclusão parcial não constitui problema.

Um modo de representação mais avançado consiste em incluir a dimensão temporal à representação, o que permite a inclusão de características relacionadas com o movimento [Chomat, 1998]. Considere-se, por exemplo, em [Kervrann, 1998] onde a representação do objecto depende da descrição das deformações globais e locais aplicadas sobre uma forma protótipo. Nesta abordagem para a modelação e aprendizagem automática dos modos de deformação de objectos não rígidos, é previamente definida uma forma protótipo (2D) sob a qual é aplicado o processo de deformação. A técnica consiste em descrever as deformações a partir dos modelos estatísticos e a estimativa *bayesiana* óptima dessas deformações é calculada utilizando técnicas de optimização. Esta modelação é adequada para a

representação de grandes classes de objectos deformáveis, podendo ser adaptada para fazer segmentação de imagens.

1.6 Seguimento

O seguimento de movimento é uma área de investigação que pode ser abordada segundo vários pontos de vista. Geralmente, o processo de seguimento envolve o emparelhamento de *pixels*, de pontos, de linhas, de áreas, de formas ou de outra informação visual [Aggarwal, 1999]. No âmbito deste trabalho, consideramos o seguimento como sendo o estabelecimento de relações coerentes com o objecto e/ou os seus componentes entre imagens, por vezes também designadas por quadros, baseado em características como posição, velocidade, forma, textura e/ou cor.

As dificuldades inerentes à tarefa de seguimento prendem-se com a complexidade da cena e dos objectos seguidos. Seguir mais pontos de um objecto, pode ser equivalente a seguir simultaneamente múltiplos objectos, uma vez que estes poderão separar-se ou fundir-se em novos objectos devido a mudanças de topologia, à oclusão, ao ruído ou às alterações de iluminação [Moeslund, 2001].

O seguimento pode ser considerado como um processo individual, ou como um meio de pré-processamento dos dados para aplicações de reconhecimento de objectos, estimação de pose, análise médica, etc. Assim, consoante o objectivo do seguimento podem ser escolhidas diferentes características a seguir. O critério de selecção das características a seguir deve prender-se também com a robustez ao ruído, às variações de intensidade, ao nível de contraste e à dimensão das imagens. As características de baixo nível, como pontos, são mais fáceis de extrair mas relativamente mais difíceis de seguir que as características de alto nível, como linhas, áreas, polígonos ou volumes 3D [Aggarwal, 1999], o que dificulta a sua análise em tempo real [J Wang, 2003]. Se, por outro lado, o seguimento constituir o pré-processamento da análise do movimento, nomeadamente no reconhecimento, então o seu objectivo principal é extrair informação específica relativa a cada imagem tanto de baixo nível, como a detecção de orlas de intensidade, ou de alto nível, como a determinação das mãos e da cabeça nas imagens.

O seguimento é usualmente suportado por predição: quando os objectos já foram previamente detectados e se tem conhecimento prévio de algumas das suas características de alto nível, o quadro seguinte pode ser previsto. A previsão introduz uma região de interesse tanto no espaço de imagem como no espaço de estados, e como tal, facilita o processamento total necessário [Moeslund, 2001]. A predição de vários parâmetros de estado baseia-se num modelo da sua evolução temporal: um modelo de velocidade e de aceleração [Gavrila, 1996] ou até de movimentos como o andar no caso do movimento humano. Uma abordagem alternativa consiste em aprender modelos probabilísticos do movimento em causa.

Neste trabalho distinguiremos as abordagens consoante sejam ou não utilizados modelos probabilísticos do movimento.

1.6.1 Estabelecimento de correspondências entre características

O trabalho pioneiro do seguimento de pessoas no mundo real foi realizado por Hogg [Hogg, 1984], onde é apresentado o *Walker* que obtém o seguimento de “baixo nível”. Para tal é utilizado um modelo do corpo humano para estimar as posturas das pessoas. A aquisição dos dados do modelo é obtida interactivamente a partir de uma sequência protótipo de imagens. Para fazer o seguimento são comparados os contornos do modelo com os pontos onde os modelos do contorno escondidos não foram removidos (Fig. 5).



Fig. 5. O modelo *Walker* [Hogg, 1984].

Em [Cai, 1995] o movimento é seguido e estimado ao estabelecer correspondências entre os segmentos de recta da imagem de fundo, realizando-se a recuperação dinâmica das imagens do fundo.

Um exemplo de seguimento de alto nível será o *Pfinder* descrito em [Wren, 1997] que utiliza um conjunto de quadros, cada um especificado em termos estatísticos por distribuições *Gaussianas* de espaço e cor, o que permite o agrupamento dos *pixels* em regiões (*blobs*) através do algoritmo da probabilidade máxima *a posteriori* fig. 6.

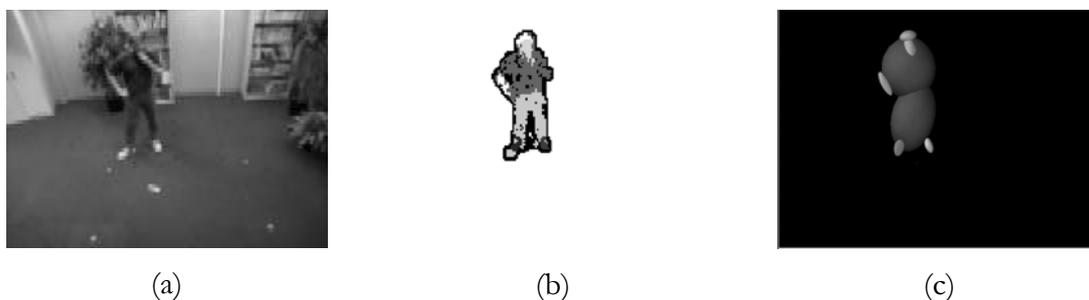


Fig. 6. Seguimento realizado pelo *Pfinder* de [Wren, 1997]:
(a) Dados obtidos por vídeo; (b) resultado da segmentação;
(c) uma representação 2D dos *blobs* estatísticos.

Outro exemplo de seguimento pode ser encontrado em [Plänkers, 1999], onde é realizada a predição linear da posição de pontos a partir do quadro anterior recorrendo ao estabelecimento de correspondências pelo método dos mínimos quadrados. Assim, é recuperada a informação relativa às formas e localização das juntas/ligações do corpo humano com intervenção mínima do utilizador.

Outro exemplo será o trabalho descrito em [Polana, 1994], no qual a informação do fluxo e da amostragem são utilizados para a representação de informação relativa à imagem de uma forma compacta, o que pode ser processado para que o classificador reconheça diferentes classes do

movimento humano, por exemplo correr ou andar. Neste trabalho, cada indivíduo é delimitado por uma caixa rectangular cujo centróide é a característica a seguir, o que permite também que o movimento continue a ser seguido mesmo quando ocorre oclusão de sujeitos. Em [Little, 1998], é utilizada uma abordagem sem modelo, à semelhança do realizado em [Polana, 1994], mas o seguimento do sujeito é obtido a partir dos momentos da distribuição de fluxo óptico denso.

Por outro lado, em [Georgescu, 2004] encontra-se um exemplo que pode ser utilizado no seguimento de características de baixo nível. Através da abordagem apresentada será possível obter o emparelhamento de pontos sob grandes deformações da imagem e alterações de intensidade de iluminação, conciliando a usual análise do fluxo óptico com um método de emparelhamento das distribuições de cor das partes consideradas relevantes nas imagens dadas.

Em [Lipton, 1998] é implementado um algoritmo de seguimento que combina o emparelhamento baseado nas correlações de aparência e a detecção do movimento, em que as regiões de movimento são utilizadas para conduzir o processo de correlação e actualização dos *templates* considerados. O trabalho mencionado permite o seguimento de movimento humano e de veículos, obtendo-se resultados satisfatórios mesmo com fundos complexos, não sendo necessário fazer previsões uma vez que o seguimento é baseado na detecção do movimento.

Os modelos de diferenciação temporal têm vindo a desenvolver-se no sentido de não utilizarem filtros de predição no seguimento robusto do movimento. Nestas metodologias, as derivadas temporais e a correspondência de orlas de intensidade podem ser combinadas para a segmentação de regiões em movimento.

1.6.2 Modelos Probabilísticos

Nesta secção abordaremos alguns modelos probabilísticos, nomeadamente o filtro de *Kalman*, e filtros de partículas.

1.6.2.1 Filtragem de Kalman e suas variantes

A filtragem de *Kalman* e suas variantes, podem ser utilizadas no seguimento, uma vez que o seu algoritmo é recursivo e portanto é uma estrutura adequada para obter a estimativa incremental do movimento. Contudo, os resultados obtidos nem sempre são os melhores, pelo que se conciliaram técnicas com o filtro de *Kalman* e desenvolveram-se algumas variantes do mesmo.

Em [Tavares, 1995] é realizado o seguimento de segmentos de recta (determinados por detecção das orlas de intensidade e simplificados por um algoritmo de faixas dinâmicas), baseado em filtragem de *Kalman* e utilizando um modelo da estrutura a seguir continuamente actualizado, obtendo-se em situações reais resultados bastante satisfatórios.

Em [Nascimento, 2003a] utiliza-se a predição de contornos através de um filtro de associação de dados probabilísticos da forma [Kirubarajan, 2004], atribuindo um grau de confiança para as características seguidas (no caso, centróides e traços - características de nível médio), o que permite que os *outliers* não

influenciem os resultados do seguimento. Desta forma, obtém-se resultados mais robustos do que através do filtro de *Kalman* (ver fig. 7).

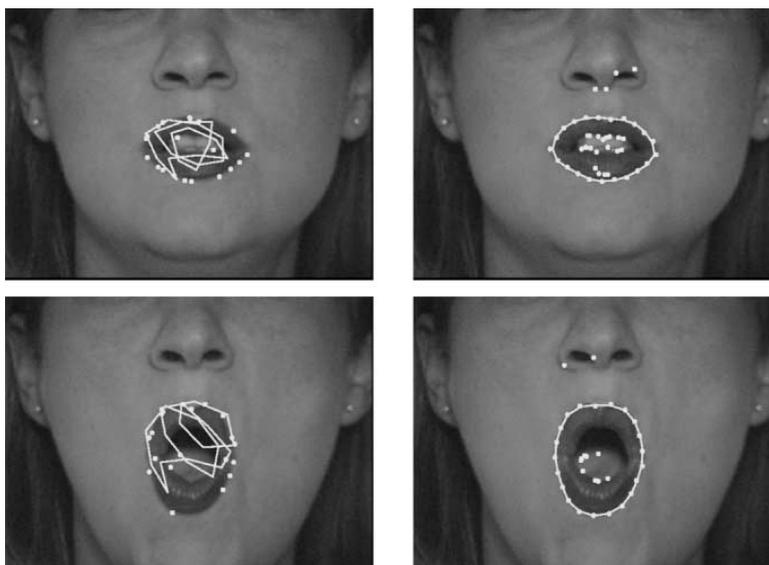


Fig. 7. Seguimento de transições abruptas com filtragem de *Kalman* (coluna da esquerda) e com uma das metodologias propostas em [Nascimento, 2003a].

O filtro de *Kalman* pressupõe a linearidade do sistema e que o ruído envolvido é do tipo branco, isto é de média nula e variância constante [Correia, 1995], estando assim restrito a situações em que a distribuição dos parâmetros de estado são unimodais. Contudo, conforme foi experimentado em [Deutscher, 1999] a natureza complexa da captura de movimento associada à existência de oclusão, fundos e dinâmicas complexas, provoca que a densidade posterior seja não *Gaussiana* e multimodal.

Para superar estas situações poderá ser mais adequado o uso de um filtro de *Kalman* estendido. Por exemplo em [Rosales, 1998], onde é proposto um mecanismo de seguimento que combina o processamento de imagem (baixo nível) com a estimativa recursiva de trajectória (alto nível) através do filtro de *Kalman* estendido, o que resulta num sistema que pode segmentar e seguir objectos em movimento antes, durante e depois da oclusão. Também em [Jang, 2002] é utilizado um filtro de *Kalman* estendido, o filtro de *Kalman* estrutural, que utiliza a informação relativa entre sub-regiões dos objectos em movimento, mas cujos resultados falham caso o modelo inicial esteja oculto.

1.6.2.2 Filtros de Partículas

Entretanto, verificou-se que os filtros de *Kalman* e suas variantes falham catastroficamente nos casos em que a função de probabilidade diferem da normal [Davison, 2001]. Assim, desenvolveram-se algoritmos alternativos para o seguimento, no sentido de se poderem considerar por exemplo hipóteses múltiplas, isto é, que suportem distribuições multimodais. O método mais reconhecido será o de *Condensação* [Isard, 1998], um filtro de partículas, que utiliza a amostragem factorizada com um modelo estocástico de movimento de objectos. Este método baseia-se na amostragem da distribuição posterior estimada do quadro anterior, e propaga essas amostras para formar o *posterior* do quadro actual. Os filtros de

partículas apresentam-se como uma alternativa poderosa ao filtro de *Kalman* [Isard, 1998], [Sidenbladh, 2003]; contudo, uma vez que são não paramétricos, requerem a utilização de um número relativamente elevado de amostras. Por exemplo, em [Davison, 2001] é preterido o uso do filtro de *Kalman* estendido para implementar um filtro de partículas, o que possibilitou o seguimento de movimentos ágeis do corpo humano.

Para problemas de maiores dimensões, poderá ser necessário utilizar métodos mais eficientes. Por exemplo, em [Cham, 1999] só os picos da distribuição anterior são amostrados e propagados ao quadro seguinte, e a distribuição é composta por *Gaussianas* parcelares cujas médias são dadas pelas amostras propagadas e as covariâncias pelas incertezas das predições, obtendo-se uma distribuição paramétrica. Em [Deutscher, 2001] é introduzido um filtro de partículas modificado que utiliza o princípio da continuação, e a combinação das orlas de intensidade e silhuetas do primeiro plano, que obtém bons resultados de seguimento [Mikic, 2003].

2. Análise de Movimento Humano

O trabalho desenvolvido no seguimento e análise de movimento pode ser utilizado no caso específico do movimento humano. Contudo, existem particularidades relativas ao movimento humano que podem ser aproveitadas no sentido de se obterem melhores resultados e em menos tempo, ou seja com menor esforço computacional.

Nesta vasta área de análise de movimento humano, existem desenvolvimentos dedicados exclusivamente às expressões faciais, como por exemplo os trabalhos descritos em [Essa, 1995] ou em [Ma, 2004], aos gestos das mãos, como os indicados em [Yin, 2001]. A título exemplificativo, a análise do movimento das mãos é importante enquanto meio de comunicação em aplicações clínicas ou não, ou em sistemas de interface homem/máquina [Pavlovic, 1997].

Neste trabalho faremos apenas referência aos estudos de movimento que podem ser estendidos à análise do movimento de corpo inteiro. Assim, no seguimento humano podem ser utilizados modelos que facilitam o processamento e se adaptam às especificidades deste tipo de movimento.

Na sequência do seguimento do movimento humano, surgem aplicações concretas específicas como são por exemplo, o reconhecimento que integra a estimação da pose, o reconhecimento do tipo de andar, etc.

2.1 Captura de Movimento

Os sistemas utilizados para capturar o movimento humano são compostos por subsistemas de detecção e processamento. A detecção pode ser feita de forma passiva ou activa, sendo que na detecção activa são colocados dispositivos nos sujeitos e nos espaços envolventes que emitem e recebem sinais, respectivamente. A detecção activa permite o processamento mais simples e é geralmente utilizada em

ambientes “controlados”. Por outro lado, a detecção passiva é baseada em fontes “naturais” de sinal, por exemplo luz visível ou outros comprimentos de onda electromagnéticos, não sendo necessária a integração de dispositivos no vestuário, por vezes com a excepção da utilização de marcadores que são colocados no sujeito de forma a facilitar a captura de movimento [Moeslund, 2001].

A visão computacional através de sistemas automáticos de captura de movimento [Lee, 2000] tem vindo a desafiar a detecção activa [Moeslund, 2001].

Independentemente da forma como é realizada a captura do movimento, a sua detecção por visão computacional pode ser resultante do movimento da câmara com o sujeito estático, ou pelo contrário o indivíduo pode mover-se e a câmara permanecer estática, ou então pode haver movimento da câmara e da pessoa em simultâneo [J Wang, 2003].

2.1.1 Estrutura de um Sistema de Processamento

Para que um sistema possa processar os dados a considerar pode ser necessário inicializá-lo, o que pode consistir por exemplo no estabelecimento de restrições ou na escolha de um modelo do indivíduo. De seguida, é feito o seguimento do movimento do sujeito, o que implica que a pessoa seja segmentada em cada imagem e que sejam estabelecidas correspondências entre dados em imagens consecutivas. Assim, uma das possíveis aplicações poderá ser a estimativa da pose do corpo do sujeito para, por exemplo, reconhecer as suas acções ou para fazer a análise clínica da sua marcha.

Um sistema de análise de movimento humano poderá não incluir todas estas fases, ainda que todos os sistemas possam ser incorporados nesta estrutura [Moeslund, 2001].

2.2 Restrições

Na análise de movimento humano podem ser consideradas duas grandes classes de restrições: as baseadas no movimento e as baseadas na aparência.

As restrições de movimento estão relacionadas com os movimentos dos indivíduos e/ou da(s) câmara(s) envolvidas; por exemplo, se a pessoa permanece dentro do espaço que esta a ser considerado na captura, se existe oclusão, se a câmara permanece sem movimento ou com movimento constante, etc.

Por outro lado, as restrições de aparência estão relacionadas com o ambiente e com os sujeitos; por exemplo, existência de iluminação constante, de fundos estáticos, posição/forma inicial dos indivíduos conhecida, etc.

A inclusão de restrições tem por objectivo a redução da complexidade do problema associado à análise do movimento.

2.3 Seguimento

Fazer o seguimento ao longo do tempo, consiste em estabelecer correspondências entre as estruturas das imagens em quadros consecutivos, ou seja ao longo de sequências de imagem, baseadas em características relacionadas com a posição, a velocidade, a textura e a cor [Aggarwal, 1999].

2.3.1 Inicialização

A inicialização inclui as operações que asseguram que o sistema principia a sua tarefa com uma interpretação correcta da cena inicial; exemplos de tais operações são: a calibração da(s) câmara(s), a adaptação do sistema às características da cena, a inicialização do modelo, etc. Por vezes o termo inicialização é substituído por pré-processamento de dados [Meyer, 1997], [Rossi, 1994].

Algumas operações de inicialização podem ser realizadas antes de principiar a captura de movimento, enquanto outras poderão estar incluídas na primeira fase do processamento de cada imagem.

Esta etapa da análise de movimento poderá ser simplificada, recorrendo-se a algumas das restrições anteriormente mencionadas.

2.3.1.1 Câmaras

No que se refere aos parâmetros da(s) câmara(s) geralmente a calibração é feita *offline*, e para as câmaras estacionárias a recalibração ocasional será usualmente suficiente. Contudo, se existirem factores que se alteram significativamente, por exemplo de disposição, é preferível realizar a calibração *online* [Azarbayejani, 1996].

A adaptação às características da cena está, geralmente, relacionada com as restrições de aparência e com as restrições dos métodos de segmentação utilizados. Usualmente, são captadas imagens de referência *offline* que serão utilizadas no processamento [Moeslund, 2001]. Todavia, também existem sistemas que através de procedimentos adaptativos calculam e actualizam as características da cena em cada instante [Haritaoglu, 1998].

2.3.1.2 Modelos

Existem duas abordagens típicas para a análise de movimento, dependendo se são utilizados modelos do sujeito *a priori* ou não. A utilização de modelos do sujeito, pode ter a desvantagem da perda de generalidade do algoritmo implementado, contudo poderá ter a vantagem de facilitar a correspondência de características entre quadros consecutivos, isto é ao longo da sequência de imagens, [Aggarwal, 1999].

Os modelos também podem ser construídos ao longo da sequência de imagens, isto é em *runtime*, tal como é realizado em [Kakadiaris, 1995]. No trabalho referido, é apresentada uma abordagem que recupera o movimento humano através da análise espácio-temporal da silhueta, sem a utilização de um modelo prévio nem a segmentação das partes do corpo.

Todavia, na construção de sistemas mais rápidos e fiáveis, geralmente é utilizado um modelo previamente definido, uma vez que deste modo se torna possível lidar com situações como, por exemplo, a perda de dados devido a situações de oclusão [Baumberg, 1995].

Quando se utiliza um modelo, a sua inicialização prende-se com a pose inicial do objecto em causa, neste caso do sujeito, e com a escolha do modelo de representação para o mesmo. Tal como em outros trabalhos para a análise de movimento humano, em [Rohr, 1997] é utilizada uma abordagem baseada num modelo para estimar a posição inicial do indivíduo, que vai sendo incrementado em cada quadro. Na maioria dos casos, o problema é reduzido ao partir-se do princípio que a pose inicial do sujeito é conhecida, ou então porque existe uma determinada pose inicial, ou ainda porque o utilizador do sistema especifica a referida pose inicial. Actualmente, poucos são os sistemas que têm uma fase de inicialização com a qual a pose inicial é determinada automaticamente [Rohr, 1997], e geralmente requerem que o sujeito efectue alguns movimentos para ser possível identificar as partes do seu corpo. Até mesmo os sistemas comerciais disponíveis para a captura de movimento, requerem que o sujeito utilize marcadores, fatos ou luvas especiais [Mikic, 2003]. Por exemplo em [Baumberg, 1995], um sistema é treinado automaticamente a partir de sequências de vídeo para aprender as restrições da forma e do movimento de objectos não-rígidos. Este sistema aprende um modelo espacio-temporal baseado em princípios físicos, que prevê as alterações na forma do sujeito ao longo do tempo. Depois o modelo construído é utilizado para fazer o seguimento em tempo real, através de técnicas lineares de filtragem optimizadas com um esquema de “*feedback*” para melhorar a robustez e precisão dos resultados. Uma abordagem alternativa é utilizada em [Mikic, 2003], assim a aquisição do modelo faz-se por localização das partes do corpo, através de um procedimento de ajustamento e crescimento que recorre a informações prévias sobre as formas e as dimensões médias das partes do corpo, sendo de seguida refinado o modelo inicial obtido através da utilização de uma rede *bayesiana*. Por outro lado, em [Ramanan, 2003] é apresentada uma abordagem de seguimento baseada num modelo de aparência robusto, já que a configuração do corpo pode mudar na sequência mas a geralmente sua aparência não, sem inicialização manual e que consegue gerir situações de oclusão temporária.

Os modelos de aparência ajustam a imagem inicial ao modelo, através da minimização do erro, podendo a forma e a aparência ser modeladas independentemente ou parametrizadas com um único conjunto de parâmetros lineares [Matthews, 2004].

Um exemplo de um modelo mais sofisticado é apresentado em [Thalmann, 1996], [Plänkers, 1999], [Plänkers, 2001], que utiliza primitivas volumétricas associadas a um esqueleto articulado, cada uma gerando um campo de potencial, sendo a pele considerada como uma isosuperfície do potencial combinado (Fig. 8). Este modelo tem a vantagem de ter um número reduzido de parâmetros e é de computação rápida; contudo para inicializar o modelo proposto, é necessário que o utilizador defina a localização de alguns pontos num par de imagens.

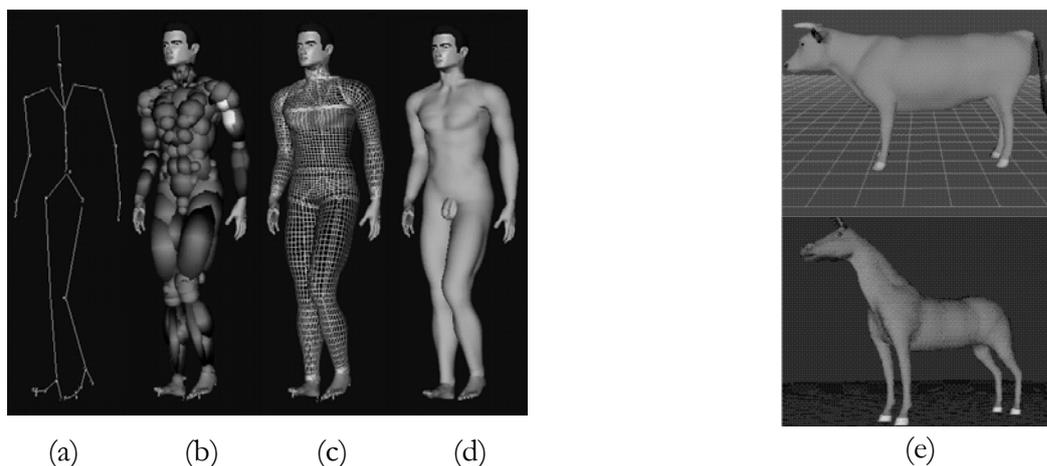


Fig. 8. Objectos articulados modelados pelo método proposto em [Plänkers, 2001]:
(a) esqueleto; (b) primitivas volumétricas utilizadas para simular músculos e tecidos gordos;
(c) superfície poligonal em representação da pele; (d) representação sombreada;
(e) uma vaca e um cavalo modelados pela mesma técnica.

Em alguns casos o algoritmo de seguimento e estimação de pose é o mesmo, quer na fase inicialização quer na fase seguinte [Njastad, 1999]. Tal implica que não sejam considerados dados temporais nem realizada a aprendizagem do sistema durante o processamento. Estes sistemas geralmente não são considerados de inicialização, uma vez que não contemplam este problema de um modo geral, mas pelo contrário estão restritos à hipótese do padrão de movimento ser conhecido.

2.3.2 Modelação

O movimento humano é complexo, não linear e varia com o tempo. Conforme foi referido, o seguimento e a análise de movimento por visão computacional pode ser realizada através da modelação do corpo humano. Os resultados de uma abordagem baseada em modelos são usualmente dependentes da adequação do modelo ao sujeito, isto é, da complexidade do modelo [Moeslund, 2001]. Alguns sistemas utilizam um modelo genérico construído a partir da média de vários indivíduos [Baumberg, 1995]. Outros sistemas consideram o sujeito actual e geram um modelo a partir dos dados observados, e podem refinar o modelo à medida que se vai obtendo mais informação acerca do sujeito em causa [Wren, 1997]; ou então pode ser obtido um modelo personalizado ao adequar um modelo genérico aos dados actuais [Zheng, 1998], por exemplo à altura do indivíduo [J Wang, 2003]. Em [Hogg, 1984] é realizada uma das primeiras aplicações de modelos articulados ao movimento do corpo humano [Little, 1998], sendo realizado o seguimento de “baixo nível” do movimento humano.

Os modelos utilizados no seguimento de movimento humano podem modelar o movimento articulado com ou sem informação prévia sobre a forma. Se apenas for considerada informação acerca da textura ou cor do indivíduo, então são utilizados modelos de aparência (fig. 9). Caso, sejam incluídas informações prévias acerca das formas, os modelos poderão ser cinemáticos, dinâmicos ou deformáveis [J Wang, 2003].

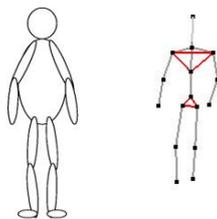


Fig. 9. Alguns exemplos de modelos utilizados para análise de movimento humano de corpo inteiro poderão ser os modelos de contornos ou os esqueletizados [J Wang, 2003].

As abordagens baseadas em modelos de aparência têm utilizado modelos *Gaussianos*, filtragem de *Kalman*, diferenciação temporal, agrupamentos de regiões ou contornos activos.

A detecção de movimento e o seu seguimento podem ser implementados como um problema de propagação para a frente, em que a diferença entre quadros pode ser modelada como a mistura de duas distribuições *Gaussianas*. Contudo, este modelo não é capaz de gerir os casos em que o fundo da cena tem textura significativa junto aos contornos dos objectos, mas é muito rápido.

A filtragem de *Kalman* e suas variantes também têm sido utilizadas no seguimento do corpo humano. Assim, para resolver problemas relacionados com o desaparecimento de dados devido a oclusão, pode ser utilizada uma versão do filtro de *Kalman* estendido que é o filtro de *Kalman* estrutural previamente mencionado [Jang, 2002]. Este filtro utiliza a informação relativa entre regiões de um objecto em movimento, mas falha caso o modelo esteja ocluso logo no instante inicial. Em [Metaxas, 1993], o seguimento é realizado utilizando uma abordagem física, e através da diferença entre a imagem prevista e a imagem actual, são calculadas as forças a aplicar ao modelo. Esta dinâmica é modelada através de um filtro de *Kalman* estendido, que se revelou um estimador sofisticado e poderoso, mas quando a função de probabilidade difere muito da normal, os resultados podem ser catastróficos [Davison, 2001]. Os modelos de diferenciação temporal têm vindo a ser desenvolvidos para superar a necessidade da utilização de um filtro predictivo, como o filtro de *Kalman*, para obter um seguimento robusto. As derivadas temporais e a correspondência entre contornos pode ser combinada para a segmentação de regiões de objectos em movimento.

O agrupamento de regiões pode ser utilizado para a segmentação de imagem, e a informação de cor para a anotação de regiões prováveis de pele do sujeito. Seguidamente, essas regiões podem ser seguidas através do seu centróide entre quadros. Em [J Gonzalez, 2003] é indicado um processo de seguimento das partes do movimento humano, que agrupa pontos da imagem com os mesmos movimentos rígidos, não sendo necessária a utilização de modelos articulados nem a sua inicialização. Já em [Kelly, 1995], são introduzidos *voxels* - conjuntos de elementos cúbicos contendo informação tal como a que objecto se refere o *pixel* em questão e a “história” desse mesmo objecto, sendo o movimento humano seguido como um grupo de *voxels* a partir do melhor ângulo de visão. Em [Mikic, 2003] a utilização de *voxels*

revelou-se como sendo uma forma robusta de seguimento, uma vez que permite grandes deslocamentos entre quadros.

A informação de cor também tem sido utilizada no seguimento de pessoas. O método de contornos não requer iluminação homogénea, mas assume o contraste significativo entre a pessoa e o fundo da cena. O seguimento de contornos utilizando contornos activos é bastante comum e tem sido evidenciado em vários trabalhos como sendo um bom método de seguimento em tempo real. Por outro lado, nos modelos que utilizam informações prévias acerca do movimento humano, podemos encontrar os modelos cinemáticos que analisam o corpo humano em termos dos seus graus de liberdade. Um exemplo poderá ser encontrado em [Yamamoto, 1991] onde são empregues modelos cinemáticos, modelos geométricos e fluxo óptico para seguir o movimento do braço e do tronco humano.

Por outro lado, quando comparando com os modelos cinemáticos, os modelos dinâmicos têm tido relativamente pouca atenção [J Wang, 2003]. A maioria dos projectos de seguimento utiliza modelos dinâmicos simples e genéricos, ou altamente específicos mas controlados manualmente. Por exemplo, em [Sidenbladh, 2003] é proposto um modelo probabilístico de aparência humana, baseado na aprendizagem de distribuições dos dados obtidos após a aplicação de filtros de imagem aos dados de treino, que pode ser utilizado mesmo em curtas sequências de imagens. Contudo, na metodologia proposta os modelos dinâmicos utilizados como são fortemente restringidos têm a sua aplicabilidade reduzida, e o sistema evolui lentamente devido ao elevado número de partículas consideradas [Davison, 2001]. De facto, as abordagens biomecânicas têm sido criticadas pelas dificuldades inerentes à medição das dinâmicas de objectos complexos que envolvem um número elevado de massas, de binários aplicados e de forças de reacção; sendo assim difícil a redução da complexidade inerente ao modelo. Uma abordagem comum consiste na utilização do filtro de *Kalman* com modelos de *Markov* não observáveis. Os modelos de *Markov* não observáveis são utilizados para determinar a forma da pessoa num quadro, e o filtro de *Kalman* utiliza os dados obtidos pelos modelos de *Markov* não observáveis para seguir a pessoa, ao estimar uma caixa que limita a trajectória e prevendo a localização da pessoa ao longo da sequência de imagens. Por exemplo, em [Marroquin, 2003] é exemplificado o uso de modelos de *Markov* não observáveis para a segmentação de imagens complexas.

Os modelos deformáveis articulados utilizam frequentemente o fluxo óptico [J Wang, 2003]. Um modelo deformável baseado em regiões pode ser construído a partir de um modelo deformável baseado em contornos, e a sua utilização conjunta poderá fazer o seguimento do movimento humano. A delimitação da região de interesse é inicializada por um algoritmo de segmentação baseada na detecção de movimento, e é seguida por um modelo deformável novo da região que explora os dados obtidos pela textura da região. O uso de um modelo deformável de regiões, baseado na textura, permite o seguimento quando existem grandes deslocamentos ou fundos complexos, e é robusto à oclusão parcial, [J Wang, 2003].

Os contornos, sendo características de nível mais elevado do que os pontos, reduzem a possibilidade de falsos emparelhamentos existente nos modelos esquelizados. O uso de contornos para representar o corpo humano, está directamente relacionado com a sua projecção nas imagens. Os modelos de contorno 2D utilizados para o seguimento humano geralmente são do tipo laço (*ribbons*) 2D. O modelo básico de laços 2D delimita as relações estruturais e de forma entre as partes do corpo, e é constituído por cinco laços com forma de “U”, sete juntas (articulações) e vários pontos médios dos segmentos. Nos laços as articulações estão localizadas entre os elementos ligados ou próximos [J Wang, 2003]. Por vezes o contorno de uma pessoa pode ser suficiente para fazer o seu seguimento; mas só funcionará correctamente, caso as poses e as perspectivas envolvidas estejam suficientemente bem representadas no conjunto de treino utilizado.

Por sua vez, os modelos volumétricos, tais como cones generalizados, cilindros elípticos e esferas, requerem um elevado número de parâmetros de modelação. Nos modelos volumétricos, uma das técnicas mais comuns é a utilização de cilindros elípticos para a modelação 3D do corpo humano [J Wang, 2003]. O modelo de cilindros, também pode ser utilizado para modelar objectos articulados que se auto-occludem, tais como os dedos humanos. Outros modelos volumétricos, assim como modelos esféricos, também podem ser utilizados; por exemplo, em [O'rourke, 1980] são utilizadas seiscentas esferas que se sobrepõem para definir um corpo humano cujo esqueleto é representado por vinte e cinco segmentos [Aggarwal, 1999].

Modelos esquelizados também podem ser utilizados para fazer a modelação do movimento humano. A representação por esqueletos, é baseada na observação de que o movimento humano é essencialmente devido ao movimento do seu esqueleto, sendo este o modo mais simples de representação. Este modelo consiste num conjunto de segmentos de recta unidos por juntas, cujo movimento permite estimar o movimento e reconhecer o comportamento de figura inteira (corpo inteiro). Este conceito foi inicialmente proposto em [Johansson, 1975] que demonstrou que os olhos humanos podem interpretar um sujeito em movimento, tal como na estrutura com dispositivos de luzes em movimento, no qual um conjunto de luzes estáticas não tem significado, enquanto o seu movimento relativo cria a impressão de uma pessoa a andar, correr, dançar, etc. Entretanto já surgiram outras variações deste modelo [J Wang, 2003]. Em geral, a estrutura do corpo é modelada por seis juntas que se encaixam numa silhueta para minimizar o ruído [J Wang, 2003]. Em [Huber, 1996] é proposta uma representação por modelos esquelizados, onde as juntas estão ligadas por segmentos de recta com um certo grau de restrição que pode ser relaxado por “molas virtuais”. A utilização destes modelos é estável em diversas circunstâncias, e o custo computacional envolvido permite o seguimento em tempo real.

A maioria das abordagens requer um modelo esquelizado do corpo, mas também poderá ser feita a combinação destes modelos com modelos volumétricos, ou primitivas 3D [J Wang, 2003]. Contudo

existem outras alternativas propostas para além das referidas; considere-se por exemplo o trabalho descrito em [Lee, 2000] no qual o corpo humano é modelado pelo Modelo de Triângulos de *Delaunay*.

Os modelos mais comuns são baseados na análise de diferentes graus de liberdade das componentes do movimento humano [Gavrila, 1996], [Kakadiaris, 1995], [Kakadiaris, 1996], [Sidenbladh, 2003]. Outras abordagens utilizam modelos hierárquicos, marcadores, e a combinação entre informações de cor e de movimento. Considere-se por exemplo, o modelo hierárquico treinado a partir de exemplos reais usando uma mistura de *Gaussianas*, para traduzir a geometria e a cinemática e modelos de *Markov* não observáveis para analisar a dinâmica envolvida [J Wang, 2003].

A conjugação de técnicas parece ser eficiente, considere-se por exemplo a proposta de [Jesus, 2002] cuja abordagem permite o seguimento humano através de um sistema interativo, com previsão múltipla, baseado em três modelos: um modelo articulado 2D, um modelo dinâmico e um modelo de aparência.

Segundo [Davison, 2001] alguns dos mais conhecidos resultados do seguimento 2D e 3D podem ser encontrados em [Bregler, 1998], onde na estimação diferencial do movimento é utilizado o produto de exponenciais conjuntamente com movimentos de torção (*twist*), o que resulta num sistema de equações lineares simples que relacionam os parâmetros do modelo do corpo humano com as medições do fluxo óptico [Mikic, 2003].

Em alguns casos a informação sensorial externa pode ser útil no seguimento do corpo humano. Assim, é usual a recolha de dados utilizando equipamento de captura com marcadores de infravermelhos colocados sobre o corpo. As coordenadas de cada marcador são então estimadas entre quadros, e seguidas ao longo da sequência de imagens; por exemplo, utilizando um rede neuronal com função de base radial como em [Lakany, 1999]. Em [Fua, 2000] segundo os autores, ao contrário dos sistemas comerciais disponíveis para seguimento de movimento humano tais como *Elite^{sm1}* e o *VICON^{sm2}*, a reconstrução 3D dos marcadores é combinada com a recuperação do movimento, usando técnicas de previsão para resolver problemas associados a possíveis ambiguidades. Também em [Park, 2004] são utilizados marcadores para seguir e simular o movimento humano, com o algoritmo “*Motion Modification*” que adapta as imagens que compõe uma base de dados (fig. 10). A utilização de marcadores, torna a metodologia robusta a oclusões durante intervalos de tempo relativamente grandes e de fácil implementação [J Wang, 2003].

2.3.3 Seguimento Múltiplo

O seguimento de várias pessoas também pode ser interessante em diversas aplicações, podendo para tal ser utilizadas uma ou várias câmaras. As abordagens utilizadas para seguir várias pessoas, de modo análogo ao seguimento de uma só pessoa, são baseadas ou na aparência (por exemplo utilizado modelos *Gaussianos*, informação temporal e de cor), ou utilizam modelos cinemáticos (por exemplo de

¹ <http://www.bts.it/en/index.html>

² <http://www.vicon.com/proddetail.jsp?id=164>

contornos), [J Wang, 2003]. Um exemplo de seguimento de várias pessoas numa mesma cena pode ser encontrado em [Davis, 1997; Marques, 2003] através da decomposição do problema em duas camadas: a detecção espacio-temporal com operações de baixo nível; e a etiquetagem dos segmentos detectados através de redes *bayesianas*. O seguimento múltiplo gere ainda problemas relacionados com a oclusão, a fusão e a separação de zonas afectas aos indivíduos. Por exemplo, em [Baumberg, 1995] e em [Mittal, 2003] essas dificuldades são superadas, sendo feito o seguimento de várias pessoas em tempo real.

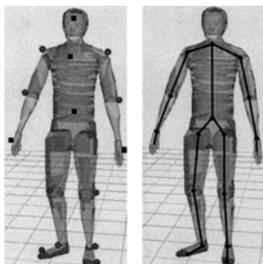


Fig. 10. Distribuição de marcadores e ligações cinemáticas resultantes [Park, 2004].

O processo de seguimento do movimento humano também pode depender do número de câmaras utilizadas. Por exemplo, o seguimento a partir de uma sequência monocular é dificultado pelo elevado número de graus de liberdade, alguns dos quais não detectados devido à ambiguidade inerente à projecção de 3D para 2D [J Gonzalez, 2003]. A título exemplificativo, considere-se o algoritmo proposto em [Han, 2004] que permite o seguimento de vários objectos, a recuperação da estrutura da cena e o movimento da única câmara utilizada. Por outro lado, nos sistemas que têm acesso a mais do que uma câmara, inicia-se o seguimento a partir de uma única câmara e quando o sistema prever que a câmara activa já não tem a visibilidade adequada do objecto seguido, o seguimento continua a partir de uma outra câmara; ver por exemplo [Kakadiaris, 1996] que utiliza três câmaras numa configuração mutuamente ortogonal. Neste caso, o sistema prevê a visibilidade do objecto através da quantificação dos pontos de vista individuais, em termos de ambiguidade, oclusão e credibilidade dos dados, o que é conseguido através da realização de várias medições [Moeslund, 2001].

Por outro lado, a maioria dos métodos que utilizam modelos 3D para fazer seguimento, geralmente têm a desvantagem de apenas lidar com movimentos reduzidos e de necessitar de imagens de elevada qualidade para realizar devidamente a segmentação [J Wang, 2003]. Alguns estudos que utilizam múltiplas câmaras superam essa desvantagem, através da modelação das forças entre o modelo 3D utilizado e os contornos na imagem da pessoa em movimento [J Wang, 2003].

Outros exemplos de seguimento com várias câmaras podem ser encontrados. Por exemplo, em [Cai, 1996] é descrito um sistema com um modelo simplificado de cabeça-tronco para seguir o movimento humano com várias câmaras, utilizando a posição e as velocidade de pontos, assim como a intensidade média dos pontos da vizinhança. A aplicação deste método tem a vantagem dos pontos serem facilmente correspondidos entre as várias câmaras por geometria epipolar; contudo, o seguimento robusto de pontos instáveis em sequências longas mantém-se difícil. Em [Mittal, 2003] é apresentado o

sistema “*M₂ Tracker*” totalmente automático capaz de segmentar, detectar e seguir várias pessoas numa cena utilizando várias câmaras de segurança sincronizadas (fig. 11).



Fig. 11. Resultados da detecção e seguimento do “*M₂ Tracker*” [Mittal, 2003]

Caso a captura de movimento seja realizada em ambientes não controlados, o número de câmaras utilizado é geralmente limitado. Também poderão acrescer problemas associados à não calibração das câmaras, dificultado pelo facto de geralmente serem portáteis, não esquecendo a possibilidade de oclusão dos vários elementos presentes na cena. Nestas circunstâncias, em [Liebowitz, 2003] é apresentado um modelo construído a partir da consideração de que os comprimentos dos segmentos do corpo se mantêm constantes ao longo do tempo. Contudo esta solução ainda depende da selecção manual das juntas, e da qualidade das imagens consideradas, geralmente obtidas em circunstâncias não controladas (fig. 12).



Fig. 12. (a) e (b) Quadros simultâneos de um salto triplo a partir de dois pontos de vista distintos; (c) Modelo considerado [Liebowitz, 2003].

2.4 Reconhecimento

O trabalho de reconhecimento de gestos, de actividades e de indivíduos é muito vasto; ver por exemplo, [Polana, 1994], [Watta, 2000], [Bobick, 2001], [Rahman, 2002], [Tominaga, 2002], [Ozer, 2002], [Yamaguchi, 2003]. Neste trabalho faremos apenas uma análise sucinta a algumas das abordagens utilizadas para reconhecer o movimento do corpo humano, como por exemplo através do emparelhamento de *templates*, do espaço de estados, de marcadores, de modelos articulados ou de modelos deformáveis [Wang, 2003].

Os diferentes movimentos humanos podem ser representados usando *templates* temporais, isto é vectores de imagens estáticas, onde o valor do vector em cada ponto é função das propriedades do movimento nas localizações espaciais correspondentes ao longo de uma sequência de imagens [Davis, 1997]. Contudo esta abordagem tem a desvantagem de ser demasiado sensível à variação da duração do movimento, ou seja ao comprimento da sequência de imagens. A abordagem do espaço de estados evita este problema ao definir cada postura estática como um estado. Desta forma, a sequência de movimento é traduzida como uma sequência de estados e a transição entre estes é definida por probabilidades [J Wang, 2003].

O reconhecimento pode também ser obtido através de marcadores, fornecendo informação vital acerca do movimento do corpo humano. Verifique-se o exemplo, mencionado previamente, do trabalho descrito em [Johansson, 1975], cujo dispositivo de luzes em movimento revelou a capacidade de reconhecer o tipo de movimento humano a partir do movimento de alguns pontos. Geralmente, os marcadores são colocados nas articulações, como por exemplo os cotovelos ou os ombros.

Os modelos articulados conseguem modelar de uma forma realista o movimento humano, e têm sido explorados para estimar a postura. Nestes modelos, cada parte constituinte é considerada rígida, mas o movimento do objecto como um todo não é admitido como rígido mas como articulado. O corpo pode ser modelado utilizando um enquadramento detalhado, e a informação extraída de cada imagem é interpretada por uma rede baseada na estrutura do modelo humano. Este modelo pode ser utilizado para prever ou antecipar futuras posições do corpo.

Os modelos deformáveis permitem a análise do movimento. Esta análise é baseada no estudo espacio-temporal da silhueta do sujeito, a partir de sequências de imagens geralmente adquiridas por várias câmaras em simultâneo. Estes métodos são usualmente baseados no uso dos contornos ocultos e dispensa a utilização de marcadores ou outros dispositivos, simplificando as dificuldades inerentes à existência de oclusões, ver por exemplo [Kakadiaris 2000].

2.4.1 Estimação da Pose

Uma das aplicações da análise de movimento humano pode ser a estimativa da postura tridimensional, por exemplo no sentido de compreender o comportamento do sujeito observado. A maioria dos estudos de estimação de pose, é baseada em abordagens que dependem da perspectiva ou de modelos geométricos. As abordagens dependentes da perspectiva armazenam um conjunto de perspectivas 2D do corpo humano em várias configurações e segundo vários pontos de vista. Caso uma perspectiva seja suficientemente próxima a um exemplo predefinido em termos de configuração e de pose, a correspondência será bem sucedida. Em alternativa, os modelos geométricos podem ser utilizados para a estimação da pose, sendo o modelo composto, por exemplo, por um conjunto de cilindros que se enquadram no perfil das partes em movimento. Os resultados experimentais demonstram que estes

modelos obtém bons resultados mesmo com fundos complexos, com condições de iluminação reduzida, e mesmo quando se verificam grandes deslocamentos [J Wang, 2003].

Em [David, 2004] é proposta uma abordagem que, simultaneamente, estima a pose e estabelece a correspondência entre objectos 3D, sem se terem determinado previamente as correspondências. Já em [Li, 2004] é proposto um algoritmo genérico para a estimação da pose, que utiliza a correspondência do tipo usual “de um para um” entre pontos de um modelo e um conjunto de pontos observados, também sem informação prévia e sem a inicialização do sistema (fig. 13).

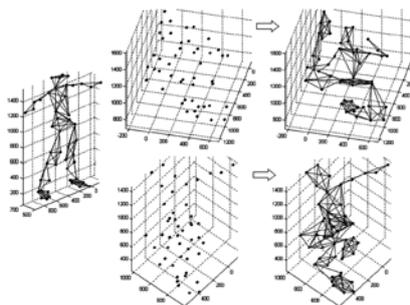


Fig. 13. Estimação de pose pelo método proposto em [Li, 2004] utilizando 47 pontos que compõe o modelo (esquerda), as poses não identificadas (ao centro) e as mesmas após identificação (direita).

2.4.2 Reconhecimento do Andar e de Gestos

Existem quatro grandes áreas de investigação na análise de movimento humano: a cinemática, a cinética, a electromiografia e a engenharia matemática, [J Wang, 2003].

A cinemática consiste na medição do movimento, e as primeiras pesquisas nesta área datam de 1870 por *Marey* e *Muybridge* [J Wang, 2003]. Os sistemas de análise do andar mais recentes, providenciam informação cinemática adicional acerca do movimento das juntas (articulações) baseada em dados cinemáticos, plataformas de forças e no uso da engenharia matemática.

A electromiografia (*EMG*) consiste na medição da actividade eléctrica dos músculos, e foi desenvolvida durante a primeira metade do séc. XX, sendo que os primeiros estudos sobre o andar com esta técnica ocorreram nas décadas de 40 e 50 por Californian [Inman, 1981].

A primeira grande aplicação da engenharia matemática ao estudo do andar, ocorreu no início dos anos 80, com uma publicação de *Braune* e *Fischer* [J Wang, 2003]. Hoje em dia, a análise do andar envolve o uso de “dinâmica inversa”, sendo que os dados são obtidos a partir do movimento dos membros e da reacção de forças medidas em diferentes dispositivos como por exemplo em plataformas de forças.

A análise do andar tem vindo a ser realizada segundo diferentes perspectivas; por exemplo, enquanto assinatura biométrica, para identificação de pessoas ou para distinguir sexos;, enquanto que em medicina, tem sido considerada para a detecção e diagnóstico de anomalias do andar ou da marcha. Com este propósito de análise do andar, é reunido um conjunto de treino de imagens, e aplicam-se sobre as mesmas algumas ferramentas estatísticas que permitem a caracterização do andar, ou guardam-

se os *templates* ou modelos para posteriormente estabelecer correspondências. As abordagens estatísticas incluem a utilização de espaços próprios, de informação espacio-temporal, de séries de tempo e de silhuetas. As abordagens baseadas em modelos, incluem o uso de modelos de *Markov* não observáveis, contornos activos, modelos esqueléticos, modelos de movimento [J Wang, 2003]. A título de exemplo, pode-se verificar em [L Wang, 2003] uma abordagem que realiza o reconhecimento do andar através da análise de silhuetas e da transformação no espaço modal baseado na análise de componentes principais. Outros exemplos podem ser encontrados em [Davis, 1997], [Davis, 1999], [Bobick, 2001]; nestes trabalhos, as abordagens apresentadas permitem o reconhecimento a partir de histogramas do movimento. Já a metodologia proposta em [Kojima, 2002] transforma as sequências de imagens representadas sob a forma de dados geométricos/numéricos em descrições textuais da acção desempenhada (fig. 14).



Uma pessoa entrou na sala.
Uma pessoa está em frente à porta.
Uma pessoa aproximou-se da mesa.

...

...



Uma pessoa está em frente à mesa.
Uma pessoa pousou um portátil na mesa.

Fig. 14. Em [Kojima, 2002] as sequências de imagens em estudo são traduzidas em texto.

No caso do movimento repetitivo, pode ser realizada a segmentação do movimento para decompor acções complexas e repetitivas numa sequência de eventos, por exemplo andar, saltar, recorrendo a modelos lineares dinâmicos simples tal como é realizado em [Lu, 2004].

Outro exemplo do trabalho de reconhecimento desenvolvido, pode encontrar-se em [Bodor, 2003] no qual, no sentido de aumentar os conjuntos de treino, são reconstruídas imagens a partir de diferentes perspectivas.

O reconhecimento tem evoluído consideravelmente [Bobick, 2001], podendo não ser necessária a recuperação de propriedades tridimensionais das pessoas em causa, ou o seguimento bidimensional dos seus membros para que o reconhecimento seja conseguido com êxito [Masoud, 2003].

3. Análise de Imagem Médica

O rápido desenvolvimento e proliferação de técnicas de processamento e análise de imagem médica estão a revolucionar a medicina [Mcinerney, 1996]. A imagem médica permite a recolha de informação potencialmente vital acerca do corpo humano geralmente de forma não invasiva. O desempenho de imagem médica já ultrapassou a simples visualização e inspecção de estruturas anatómicas, tendo sido

utilizada, por exemplo, para planear e simular cirurgias [Bro-Nielsen, 1996], planear radioterapia, realizar o seguimento do progresso de doenças, etc. Assim, por exemplo, ao recolher a forma e organização detalhada das estruturas anatómicas, o cirurgião pode planear uma abordagem óptima à estrutura objectivo. Do mesmo modo, em radioterapia, a imagem médica pode permitir o cálculo da dosagem de radiação a aplicar sobre um dado tumor, com os mais reduzidos efeitos colaterais sobre os tecidos saudáveis [Mcinerney, 1996].

As aplicações de análise de imagem médica são inúmeras. Por exemplo, em [Claridge, 2003] através de análise de imagem médica, é proposto um modelo baseado em princípios físicos da coloração dos tecidos que providencia o cruzamento de dados entre as cores da imagem e os parâmetros histológicos subjacentes. Com este modelo representativo de todas as cores normais de pele, independentemente da origem racial, idade ou género, as cores anormais da pele, devido por exemplo a melanomas malignos, podem assim ser detectadas uma vez que não são conformes com o modelo.

Outro exemplo pode ser encontrado em [Chabanas, 2003], onde é apresentado um modelo 3D por elementos finitos dos tecidos flácidos da face, que pode ser utilizado em cirurgia maxilo-facial para prever as deformações desses tecidos resultantes da reposição óssea. O referido modelo adapta-se automaticamente à morfologia do paciente através da pele e das superfícies cranianas segmentadas a partir de imagens médicas. De acordo com os autores, as modificações previstas são coerentes com os resultados previstos. Em [Hilger, 2003] é modelado a forma temporal e o tamanho das mandíbulas humanas para efeitos de análise, de simulação e de predição que ajudam à compreensão do crescimento craniofacial.

Em [Cocosco, 2003], é proposto um método de classificação de tecidos cerebrais representados em imagens obtidas por ressonância magnética. O método é inteiramente automático e robusto relativamente a variabilidades anatómicas e patológicas. Este trabalho auxilia a pesquisa e os estudos clínicos de cérebros humanos, permitindo a análise quantitativa do volume de tecido em populações saudáveis ou não. Em [Paling, 2004] é exemplificado o uso de uma técnica de análise de imagem médica para estudar as atrofia cerebrais. A técnica de análise utilizada mostra-se robusta e precisa, podendo constituir uma ferramenta valiosa em futuros estudos de demência e perturbações neuro-degenerativas relacionadas.

Contudo, ainda que os dispositivos de imagem médica providenciem imagens excepcionais da anatomia humana, o uso de computadores para quantificar e analisar as estruturas representadas com exactidão e eficiência é ainda algo limitado [Mcinerney, 1996]. Para apoiar as investigações biomédicas e actividades clínicas do diagnóstico à cirurgia, é necessário extrair de forma eficiente dados precisos, quantitativos e repetitivos.

3.1 Segmentação em Imagem Médica

A segmentação de estruturas anatómicas – a classificação de conjunto original de *pixels* da imagem em subconjuntos correspondentes a essas estruturas – é uma das primeiras tarefas essenciais na análise de imagem médica. Contudo, segmentar estruturas de imagens médicas e reconstruir representações compactas dessas estruturas, é um processo difícil devido ao reduzido tamanho das bases de dados existentes e à complexidade e variabilidade das formas anatómicas de interesse. Além destas dificuldades podem ainda surgir outros inconvenientes como a existência de ruído. À semelhança da análise de imagem genérica, na análise de imagem médica o principal desafio consiste em extrair elementos de fronteira que pertencem à mesma estrutura e eventualmente integra-los adequadamente num modelo consistente e coerente para a estrutura. Um exemplo primário será a segmentação do coração, especialmente do ventrículo esquerdo, que constitui um pré-requisito para obter informações como a razão de volume ventricular ou o *output* cardíaco.

A maioria da segmentação clínica é efectuada por edição manual [Mcinerney, 1996], como por exemplo a medição da forma e de volumes 3D através da análise individual de cada corte (*slice*), o que levanta muitas contrariedades como a dificuldade de obter resultados que possam ser reproduzidos ou própria a fadiga do operador. A automatização das abordagens manuais sem a utilização de modelos é difícil pela complexidade e variabilidade da forma das estruturas entre indivíduos. Assim, devido ao problema da segmentação usualmente apresentar poucas restrições, fica comprometida a eficácia de abordagens que consideram apenas informação do tipo local. Para superar estas dificuldades, geralmente são utilizados modelos deformáveis que podem contemplar informação prévia sobre a forma do objecto e assim restringir, ou seja simplificar, o problema de segmentação [Mcinerney, 1996]. Estes modelos geométricos, contínuos e interligados, consideram a fronteira de um objecto como um todo, e a sua continuidade e suavidade pode compensar o ruído, fendas e outras irregularidades presentes nas fronteiras dos objectos, [Tavares, 2000].

Os modelos deformáveis são uma das técnicas capazes de assimilar a variabilidade de estruturas biológicas ao longo do tempo, e entre vários indivíduos, e como tal são muito utilizados. Por exemplo, os modelos deformáveis podem ser empregues em imagens geradas por *raio-X*, por tomografia computadorizada, por ressonância magnética ou por ultrasons. Quer sejam 2D ou 3D, os modelos deformáveis têm sido utilizados para segmentar, visualizar, seguir e quantificar uma grande variedade de estruturas anatómicas, que em escala podem variar entre macroscópicas e microscópicas. Estas incluem o cérebro, o coração, as artérias coronárias, o fígado, os pulmões, o estômago, os rins, objectos como tumores cerebrais, fetos, estruturas celulares como neurónios ou cromossomas, ou até partes de células, como em [Bamford, 1998] onde são segmentados núcleos de células. Os modelos deformáveis têm também sido utilizados, por exemplo, no seguimento do movimento não rígido do coração ou no movimento de eritrócitos, na localização de estruturas do cérebro e no registo de imagens da retina [Mcinerney, 1996].

A utilização de modelos deformáveis em análise de imagem médica foi inicializada através da aplicação de modelos de contorno deformável, tais como *snakes* [Kass, 1988] para segmentar estruturas em imagens, ver por exemplo [Cohen, 1993], [Maurincomme, 1993], [Kapur, 1996], [Fishman, 1996], tendo evoluído com a utilização de modelos de superfícies deformáveis, ver por exemplo [Nastar, 1996].

Usualmente o modelo deformável é inicializado próximo do objecto que se pretende segmentar. Para tal, utilizam-se as capacidades interactivas destes modelos ao afiná-los, geralmente, manualmente (para verificar um estudo detalhado sobre a interacção na segmentação de imagens médicas ver [Olabarriaga, 2001]). Quando o resultado obtido numa imagem inicial for satisfatório, poder-se-á utilizar o modelo ajustado como a aproximação inicial da imagem seguinte da sequência [Tavares, 2000]. Um exemplo desta metodologia pode ser encontrado em [Bruijne, 2004], onde é realizada a segmentação interactiva de aneurismas aórticos abdominais.

Alguns investigadores têm incorporado informação acerca da forma do objecto nos modelos deformáveis através da utilização de protótipos deformáveis, tendo sido implementadas com sucesso em muitas aplicações de interpretação automática de imagens [Mcinerney, 1996]. Uns dos exemplos de protótipo utilizado são as superquadráticas, ver por exemplo [Metaxas, 1993], em que o modelo se deforma global e localmente através da utilização das equações de *Lagrange* e filtragem de *Kalman*.

Ainda que amplamente utilizadas as *snakes* e outros modelos deformáveis similares, apresentam algumas limitações na extracção de regiões de interesse em imagem médica. Por exemplo, em aplicações não interactivas, os modelos deformáveis devem ser inicializados próximos da estrutura no sentido de se garantir um bom desempenho; ou, no caso das *snakes*, as restrições de energia podem limitar a flexibilidade geométrica e assim impedir protrusões significativas. Também a topologia da estrutura a segmentar deve ser conhecida e constante, uma vez que os modelos de contorno clássicos são paramétricos e sem mecanismos adicionais são incapazes de transformações topológicas [Tavares, 2000].

Para melhorar e automatizar a segmentação via contornos deformáveis têm vindo a ser propostos vários métodos [Tavares, 2000]. Por exemplo, em [Cohen, 1991] é utilizada uma força interna de inflação que diminui a sensibilidade do contorno deformável à inicialização e ao ruído. Com o mesmo propósito em [Herlin, 1992] é proposta a incorporação de informação regional. Já em [Grzeszczuk, 1997] é utilizado um algoritmo para obter soluções globais de minimização da energia, que permite a consideração de restrições não diferenciáveis. Com o mesmo propósito, em [Delagnes, 1995] é proposta a utilização de contornos activos poligonais no caso de segmentação de imagens com fundos complexos, e que permite a aproximação de objectos mesmo constituídos por um número elevado de vértices. Em [Sebastian, 2003] são conciliadas várias metodologias numa única abordagem para fazer a segmentação e registo ao longo do tempo dos ossos do carpo. No referido trabalho, são combinadas as vantagens dos modelos de contornos activos com o crescimento de regiões e métodos de competição local e global; através da implementação de uma evolução de curva numa região em crescimento a

partir de alguns pontos, sendo este crescimento controlado por uma competição entre regiões vizinhas. Assim podem ser superadas algumas das dificuldades resultantes das não uniformidade do tecido ósseo, da irregularidade da forma de alguns ossos, dos relativamente reduzidos espaços entre ossos quando comparados com a resolução das imagens captadas, assim como da presença de vários vasos sanguíneos.

Em [Keserci, 2002] é apresentado um esquema de diagnóstico computadorizado para a detecção automática de nódulos pulmonares em radiografias digitais do peito. O esquema proposto sugere a utilização em paralelo de dois processos: Um dos processos será a filtragem adaptativa para realçar os nódulos e suprimir as estruturas pulmonares normais, seguido da convencional extração de características morfológicas. O outro processo consiste na utilização de contornos activos baseados em ondinhas (*wavelet snakes*) no sentido de aproximar a fronteira dos nódulos. Finalmente, é realizada a combinação dos resultados obtidos por cada um dos processos através da utilização de uma rede neuronal. Com este esquema faz-se a redução eficiente dos falsos-positivos após a selecção inicial de candidatos a nódulos.

Em [Mcinerney, 2002] é introduzida uma abordagem para análise de imagem médica que combina as metodologias de modelos deformáveis com conceitos de inteligência artificial. Em particular, são propostos organismos artificiais deformáveis; isto é, agentes deformáveis cuja função é a de automaticamente segmentar, etiquetar e analisar qualitativamente estruturas anatómicas em imagens médicas. Esses organismos possuem corpos deformáveis com sensores distribuídos, assim como cérebros rudimentares, e centros de percepção, de comportamento e cognitivos. Os seus comportamentos, que se manifestam voluntariamente através do movimento e alteração da forma do seu corpo, são baseados nas características das imagens, nos conhecimentos anatómicos pré-armazenados e num plano cognitivo deliberado. Com a referida técnica, é ultrapassado o ruído, a existência de orlas incompletas, as variações anatómicas consideráveis e a interferência entre estruturas colaterais.

Outra solução pode ser a incorporação prévia de informação sobre os objectos em causa nos modelos deformáveis segundo uma abordagem probabilística. Por exemplo, em [Cootes, 1994] é apresentada uma abordagem estatística para segmentar vários órgãos a partir de imagens médicas 2D e 3D. Também em [Chesnaud, 1999] é utilizada uma abordagem estatística para segmentar regiões. Outro exemplo poderá ser encontrado em [Shen, 2004], onde é apresentada uma metodologia para a construção de um atlas estatístico da distribuição espacial do cancro da próstata, que foi optimizado para as biopsias com agulha. Na referida metodologia, é utilizado um modelo deformável adaptativo para a normalização espacial e registo de amostras. Baseado no atlas construído, foi desenvolvido um modelo predictivo estatístico para maximizar a probabilidade de detecção de cancro da próstata quando é realizada a biopsia por agulhas. Também em [Rapantzikos, 2003] é utilizado um histograma equalizador adaptativo modificado para extrair informação útil de imagens da retina, no sentido de

detectar e quantificar as anormalidades da retina, como depósitos esbranquiçados/amarelados, que antecedem doenças degenerativas da mesma. Já em [Trattner, 2004], é apresentado o desenvolvimento de uma ferramenta automática para a identificação de tipos de dados microbiológicos recorrendo à visão computacional e a técnicas de modelação estatística. Com o trabalho desenvolvido, pretendeu-se evitar a análise humana dos resultados da infecção de bactérias por vírus, associada a um elevado dispêndio de tempo e sujeita a vários erros, especialmente quando a tecnologia está a permitir o considerável aumento do número de testes a realizar.

Na análise de imagem médica podem ser utilizadas outras técnicas além das anteriormente indicadas. Por exemplo, em [Heneghan, 2002] é proposta a segmentação das estruturas vasculares da retina através de técnicas de análise de imagem, como o pré-processamento morfológico, derivação, filtragem morfológica e binarização. A abordagem proposta, possibilita a identificação de pontos de cruzamento de vasos sanguíneos, cálculo da sua espessura e índice de tortuosidade. A segmentação das referidas estruturas é útil na detecção de doenças da retina, uma vez que algumas dessas doenças são caracterizadas pela alteração dos vasos retiniais; por exemplo, pela tortuosidade ou dilatação excessiva. Também em [Kondo, 2004] são utilizadas técnicas de processamento de imagem para obter a segmentação de dentes, a partir de representações digitalizadas de um modelo de estudo da dentição, para posteriormente a informação obtida ser utilizada em procedimentos ortodônticos.

3.2 Análise e Seguimento em Imagem Médica

Para realizar a análise e seguimento de objectos em imagem médica é também usual a utilização de modelos deformáveis [Mcinerney, 1996]. Por exemplo, os modelos deformáveis são utilizados para seguir estruturas não rígidas em movimento, tanto a nível microscópico como a nível macroscópico, tais como células sanguíneas [Leymarie, 1993]. Contudo, a utilização primária de modelos deformáveis para o seguimento em imagem médica tem sido na medição do comportamento dinâmico do coração humano, especialmente o ventrículo esquerdo, veja-se por exemplo [Pentland, 1991] ou [Nastar, 1994]. A caracterização das regiões da parede do coração é necessária para determinar a severidade e extensão de doenças como a isquemia [Mcinerney, 1996]. A obtenção de imagens 3D do coração ao longo do tempo é actualmente possível através de ressonância magnética e outras técnicas de aquisição de imagem, sendo a resolução espacial e temporal bastante satisfatória [Tavares, 2000].

Nas abordagens mais simples, é geralmente utilizado um modelo de contorno 2D deformável para segmentar a fronteira do ventrículo esquerdo em cada fatia (*slice*) de uma imagem inicial 3D [Tavares, 2000]. Esses contornos são depois utilizados como aproximação inicial das fronteiras do ventrículo esquerdo nas correspondentes fatias da imagem 3D no instante seguinte. Seguidamente, esses contornos são deformados de modo a extrair o novo conjunto de fronteiras do ventrículo esquerdo. Esta abordagem é utilizada por exemplo em [Herlin, 1992]. A propagação temporal dos contornos

deformáveis diminui drasticamente o tempo de segmentação manual do ventrículo esquerdo a partir de imagens 3D obtidas ao longo de um ciclo cardíaco [Tavares, 2000].

A extensão do trabalho desenvolvido na área da análise de imagem médica é comprovada pela diversidade de abordagens utilizadas. Por exemplo, em [Chen, 1998] é utilizado um modelo hierárquico do movimento do ventrículo esquerdo, construído por uma combinação de uma superquádrica globalmente deformável com uma superfície localmente deformável. Com este modelo é produzida uma decomposição hierárquica que caracteriza o movimento do ventrículo esquerdo, cuja resolução pode variar entre grosseira e fina [Tavares, 2000]. Outros exemplos podem ser encontrados, por exemplo, em [Nastar, 1996], onde são utilizados modelos deformáveis para seguir e recuperar o movimento do ventrículo esquerdo, e análise modal para parametrizar os modelos. Essa parametrização é feita a partir dos modos de vibração em regime livre, sendo o detalhe obtido determinado pelo número de modos utilizados. Já em [Shechter, 2004], é apresentado um método para medir o movimento respiratório natural do coração a partir de angiogramas coronários. Para tal é utilizado um modelo paramétrico da respiração cardíaca que decompõe o campo de deslocamentos nas suas componentes cardíaca e respiratória.

Como já foi anteriormente referido, existem algumas limitações associadas à utilização de modelos deformáveis, tais como a modelização separada das superfícies endocárdica e epicárdica [Tavares, 2000]. No sentido de superar essa dificuldade em [Park, 1996] é desenvolvido um modelo que considera a natureza volumétrica da parede do coração e que incorpora a parametrização descritiva directamente na sua formulação [Tavares, 2000]. Também nesse âmbito, em [Donnell, 1995] é utilizado um modelo híbrido e volumétrico para analisar e comparar o ventrículo esquerdo [Tavares, 2000].

Conforme mencionado anteriormente, a análise de imagem médica tem diversas aplicações que não se limitam à análise do movimento do ventrículo esquerdo. Por exemplo em [Kita, 1996] é realizada a análise e seguimento de imagens do estômago obtidas por vários pontos de vista a partir de raios-X. A abordagem utilizada baseia-se em modelos deformáveis físicos, sendo extraídas as regiões de um objecto deformável a partir de imagens obtidas segundo diferentes perspectivas.

Bibliografia

- [Aggarwal, 1999] - J. Aggarwal, Q. Cai, *Human Motion Analysis: A Review*, Computer Vision and Image Understanding 73(3), 428/440, 1999
- [Azarbayejani, 1996] - A. Azarbayejani, C. Wren, A. Pentland, *Real-Time 3D Tracking of the Human Body*, IMAGE'COM 96, Bordeaux, France, 1996
- [Bamford, 1998] - P. Bamford, B. Lovell, *Unsupervised Cell Nucleus Segmentation with Active Contours*, Signal Processing 71, 203/213, 1998
- [Baumberg, 1995] - A. Baumberg, PHD Thesis, *Learning Deformable Models for Tracking Human Motion*, School of Computer Studies, The University of Leeds, UK, 1995
- [Bobick, 2001] - A. Bobick, J. Davis, *The Recognition of Human Movement using Temporal Templates*, IEEE Transactions on Pattern Analysis and Machine Intelligence 23(3), 2001
- [Bodor, 2003] - R. Bodor, B. Jackson, O. Masoud, N. Papanikolopoulos, *Image-Based Reconstruction for View-Independent Human Motion Recognition*, Proceedings of the IEEE/RJS International Conference on Intelligent Robots and Systems, 2003
- [Bregler, 1997] - C. Bregler, *Learning and Recognizing Human Dynamics in Video Sequences*, Proceedings IEEE Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 1997
- [Bregler, 1998] - C. Bregler, J. Malik, *Tracking People with Twists and Exponential Maps*, International Conference on Computer Vision and Pattern Recognition, 1998
- [Bro-Nielsen, 1996] - M. Bro-Nielsen, *Surgery Simulation using Fast Finite Elements*, VBC'96 – 4th International Conference in Visualization in Biomedical Computing, HAMBURG, GERMANY, 1996
- [Bruijne, 2004] - M. Bruijne, B. Ginneken, M. Viergever, W. Niessen, *Interactive Segmentation of Abdominal Aortic Aneurysms in CTA Images*, Medical Image Analysis 8(2), 127/138, 2004
- [Cai, 1995] - Q. Cai, A. Mitiche, J. Aggarwal, *Tracking Human Motion in an Indoor Environment*, Proceedings 2nd. Intl. Conference on Image Processing, Washington D.C., 1995
- [Cai, 1996] - Q. Cai, J. Aggarwal, *Tracking Human Motion using Multiple Cameras*, Proceedings of International Conference on Pattern Recognition, Viena, 1996
- [Campbell, 1995] - L. Campbell, A. Bobick, *Recognition of Human Body Motion Using Phase Space Constraints*, International Conference on Computer Vision, Cambridge, Massachusetts, 1995
- [Chabanas, 2003] - M. Chabanas, V. Luboz, Y. Payan, *Patient Specific Finite Element Model of the Face Soft Tissues for Computer-Assisted Maxillofacial Surgery*, Medical Image Analysis 7(2), 131/151, 2003
- [Cham, 1999] - T. Cham, J. Rehg, *Multiple Hypothesis Approach to Figure Tracking*, Conference on Computer Vision and Pattern Recognition, Fort Collins, USA, 1999

- [Chen, 1998] - C. W. Chen, T.S. Huang, M. Arrott, *Modeling, Analysis, and Visualization of Left Ventricle Shape and Motion by Hierarchical Decomposition*, IEEE Deformable Models In Medical Image Analysis, 1998
- [Chesnaud, 1999] - C. Chesnaud, P. Réfrégier, V. Boulet, *Statistical Region Snake-Based Segmentation Adapted to Different Physical Noise Model*, IEEE Transactions on Pattern Analysis and Machine Intelligence 21(11), 1999
- [Chomat, 1998] - O. Chomat, J. Crowley, *Recognizing Motion using Local Appearance*, International Symposium on Intelligent Robotic Systems, University of Edinburgh, 1998
- [Christensen, 1997] - C. Christensen, S. Corneliussen, *Visualization of Human Motion using Model-based Vision*, Denmark, Laboratory of Image Analysis, Aalborg University, 1997
- [Claridge, 2003] - E. Claridge, S. Cotton, P. Hall, M. Moncrieff, *From Colour to Tissue Histology: Physics-based Interpretation of Images of Pigmented Skin Lesions*, Medical Image Analysis 7(4), 489/502, 2003
- [Cocosco, 2003] - C. Cocosco, A. Zijdenbos, A. Evans, *A Fully Automatic and Robust Brain MRI Tissue Classification Method*, Medical Image Analysis 7, 513/527, 2003
- [Cohen, 1991] - L. Cohen, *Note on Active Contour Models and Balloons*, CVGIP: Image Understanding 53(2), 211/218, 1991
- [Cohen, 1993] - L. Cohen, I. Cohen, *Finite Elements for Active Contour Models and Ballons for 2-D and 3-D Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1131/1147, 1993
- [Cootes, 1992] - T. Cootes, C. Taylor, *Active Shape Models: Smart Snakes*, British Machine Vision Conference, 1992
- [Cootes, 1994] - T. Cootes, A. Hill, C. Taylor, J. Haslam, *The use of active shape models for locating structures in Medical Images*, Image and Vision Computing 12(6), 355/366, 1994
- [Cootes, 1995] - T. Cootes, C. Taylor, D. Cooper, J. Graham, *Active Shape Models - Their Training and Applications*, Computer Vision Image Understanding 61, 38/59, 1995
- [Correia, 1995] - M. Correia, Dissertação de Mestrado, *Análise de Movimento em Sequências de Imagens*, Faculdade de Engenharia, Universidade do Porto, 1995
- [Correia, 2001] - M. Correia, Tese de Doutoramento, *Técnicas Computacionais na Percepção Visual do Movimento*, Faculdade de Engenharia, Universidade do Porto, 2001
- [Cucchiara, 2000] - R. Cucchiara, C. Grana, M. Piccardi, A. Prati, *Statistic and Knowledge-based Moving Object Detection in Traffic Scenes*, Proceedings of the 3rd IEEE Conference on Intelligent Transportation Systems (ITSC2000), Dearborn, Indiana, USA, 2000
- [D Lee, 2000] - D. Lee, S. Jung, *An approach to Image-Based Motion Retargeting*, Avatars 2000, Lausanne, Switzerland, 2000

- [D Shen, 2004] - D. Shen, Z. Lao, J. Zeng, W. Zhang, I. Sesterhenn, L. Sun, J. Moul, E. Herskovits, G. Fichtinger, C. Davatzikos, *Optimized Prostate Biopsy via a Statistical Atlas of Cancer Spatial Distribution*, Medical Image Analysis 8, 139/150, 2004
- [Darrell, 1994] - T. Darrell, P. Maes, B. Baumberg, A. Pentland, *A novel Environment for Situated Vision and Behaviour*, Workshop for Visual Behaviours at CVPR-94, 1994
- [David, 2004] - P. David, D. Dementhon, R. Duraiswami, H. Samet, *SoftPOSIT: Simultaneous Pose and Correspondence Determination*, International Journal of Computer Vision 59(3), 259 /284, 2004
- [Davis, 1997] - J. Davis, A. Bobick, *The Representation and Recognition of Human Movement using Temporal Templates*, CVPR, 1997
- [Davis, 1999] - J. Davis, G. Bradski, *Real-time Motion Template Gradients using Intel CVLib*, IEEE ICCV Workshop on Frame-rate Vision, 1999
- [Davison, 2001] - A. Davison, J. Deutscher, I. Reid, *Markerless motion Capture of Complex Full-Body Movement for Character Animation*, Manchester, 2001
- [Delagnes, 1995] - P. Delagnes, J. Benois, D. Barba, *Active Contours Approach to Object Tracking in Image Sequences with Complex Background*, Pattern Recognition Letters 16, 171/178, 1995
- [Deutscher, 1999] - J. Deutscher, A. Blake, B. North, B. Bascle, *Tracking through Singularities and Discontinuities by Random Sampling*, 7th. Conference on Computer Vision, 1999
- [Deutscher, 2001] - J. Deutscher, A. Davison, I. Reid, *Automatic Partitioning of High Dimensional Search Spaces associated with Articulated Body Motion Capture*, Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Kauai, 2001
- [Donnell, 1995] - T. Donnell, A. Grupta, T. Boult, *The Hybrid Volumetric Ventriculoid: A Model For Mr-Spamm 3-D Analysis*, IEEE Deformable Models In Medical Image Analysis, 1995
- [Essa, 1995] - I. Essa, A. Pentland, *Facial Expression Recognition using a Dynamic Model and Motion Energy*, Proceedings of the Fifth International Conference on Computer Vision, 1995
- [Feldman, 2003] - A. Feldman, T. Balch, *Automatic Identification of Bee Movement Using Human Trainable Models of Behaviour*, Mathematics and Algorithms of Social Insects, 2003
- [Fishman, 1996] - E. Fishman, B. Kuszyk, L. Gao, B. Cabral, *Surgical Planning for Liver Resection*, IEEE Transactions on Visualization and Computer Graphics 29(1), 1996
- [Franke, 1998] - U. Franke, D. Gavrila, S. Görzig, F. Lindner, F. Paetzold, C. Wöhler, *Autonomous Driving Goes Downtown*, IEEE Intelligent Systems 13(6), 40/48, 1998

- [Fua, 2000] - P. Fua, L. Herda, R. Plaenkers, R. Boulic, *Human Shape and Motion Recovery Using Animation Models*, 19th Congress, International Society for Photogrammetry and Remote Sensing, Amsterdam, Netherlands, 2000
- [Garcés, 1999] - D. Garcés, *Body Silhouette Extraction from Video Sequences*, Computer Graphics Laboratory (LIG), Lausanne, Switzerland, 1999
- [Gavrila, 1996] - D. Gavrila, L. Davis, *3D Model-Based Tracking of Humans in Action: A multi-view Approach*, IEEE Conference on Computer Vision and Pattern Recognition,, San Fransisco, USA, 1996
- [Gavrila, 1999] - D. Gavrila, *The Visual Analysis of Human Movement: A Survey*, Computer Vision and Image Understanding 73(1), 1999
- [Georgescu, 2004] - B. Georgescu, P. Meer, *Point Matching under Large Image Deformations and Illumination Changes*, IEEE Transactions on Pattern Analysis and Machine Intelligence 26(6), 2004
- [Gleicher, 1999] - M. Gleicher, *Animation From Observation: Motion Capture and Motion Editing*, ACM SIGGRAPH Applications of Computer Vision to Computer Graphics 33(4), 1999
- [Goncalves, 1998] - L. Goncalves, E. Bernardo, P. Perona, *Reach out and touch Space (Motion Learning)*, International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998
- [Grzeszczuk, 1997] - R. Grzeszczuk, D. Levin, *Brownian Strings: Segmenting Images With Stochastically Deformable Contours*, IEEE Transactions On Pattern Analysis And Machine Intelligence 19(10), 1997
- [Gu, 1994] - H. Gu, Y. Shirai, M. Asada, *MDL-Based Spatiotemporal Segmentation from Motion in a Long Image Sequence*, Computer Vision and Pattern Recognition, 1994
- [Han, 2004] - M. Han, T. Kanade, *Reconstruction of a Scene with Multiple Linearly Moving Objects*, International Journal of Computer Vision 59(3), 2004
- [Hanek, 2004] - R. Hanek, M. Beetz, *The Contracting Curve Density Algorithm: Fitting Parametric Curve Models to Image using Local Self-Adapting Separation Criteria*, International Journal of Computer Vision 59(3), 233/258, 2004
- [Haritaoglu, 1998] - I. Haritaoglu, D. Harwood, L. Davis, *W4: Who? When? Where? What? - A real Time System for Detecting and Tracking People*, International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998
- [Heneghan, 2002] - C. Heneghan, J. Flynn, M. O'keefe, M. Cahill, *Characterization of Changes in Blood Vessel Width and Tortuosity in Retinopathy of Prematurity using Image Analysis*, Medical Image Analysis 6(4), 407/429, 2002

- [Herlin, 1992] - I. Herlin, N. Ayache, *Features Extraction and Analysis Methods for Sequences of Ultrasound Images*, IEEE Deformable Models In Medical Image Analysis, 1992
- [Hilger, 2003] - K. Hilger, R. Larsen, M. Wrobel, *Growth Modeling of Human Mandibles using Non-Euclidean Metrics*, Medical Image Analysis 7(4), 425/433, 2003
- [Hogg, 1984] - D. Hogg, PhD Thesis, *Interpreting Images of a Known Moving Object*, University of Sussex, 1984
- [Huber, 1996] - B. Huber, *3D real-time gesture recognition using proximity space*, Proceedings International Conference on Pattern Recognition, Vienna, Austria, 1996
- [Inman, 1981] - V. Inman, H. Ralston, F. Todd, *Human walking*, 1981
- [Isard, 1998] - M. Isard, A. Blake, *Condensation - Conditional Density Propagation of Visual Tracking*, International Journal on Computer Vision, 5/28, 1998
- [Iwai, 1999] - Y. Iwai, K. Ogaki, M. Yacida, *Posture Estimation using Structure and Motion Models*, International Conference on Computer Vision, Corfu, Greece, 1999
- [Iwasawa, 1997] - S. Iwasawa, K. Ebihara, J. Ohya, S. Morishima, *Real-Time Estimation of Human Body Posture from Monocular Thermal Images*, Conference on Computer Vision and Pattern Recognition, 1997
- [J Gonzalez, 2003] - J. Gonzalez, I. Lim, P. Fua, D. Thalmann, *Robust Tracking and Segmentation of Human Motion in an Image Sequence*, ICASSP 2003 Conference on Acoustics, Speech and Signal Processing, Hong Kong, 2003
- [J Wang, 2003] - J. Wang, S. Singh, *Video Analysis of Human Dynamics - A Survey*, Real-time Imaging Journal 9(5), 320/345, 2003
- [Jain, 1998] - A. Jain, Y. Zhong, M. Dubuisson-Jolly, *Deformable Template Models: A Review*, Signal Processing 71, 109/129, 1998
- [Jang, 2002] - D. Jang, S. Jang, H. Choi, *2D human body tracking with structural Kalman filter*, Pattern Recognition 35(10), 2041/2050, 2002
- [Jesus, 2002] - R. Jesus, A. Abrantes, J. Marques, *Tracking of Human Body using Multiple Predictors*, AMDO - Articulated Models and Deformable Objects, Spain, 2002
- [Johansson, 1975] - G. Johansson, *Visual motion perception*, Science American 232(6), 76/88, 1975
- [Kakadiaris, 1995] - I. Kakadiaris, D. Metaxas, *3D Human Body Model Acquisition From Multiple Views*, 5th International Conference in Computer Vision, Boston, MA, 1995
- [Kakadiaris, 1996] - I. Kakadiaris, D. Metaxas, *Model-based estimation of 3D Human Motion with Occlusion based on Multi-viewpoint Selection*, Proceedings IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 1996
- [Kakadiaris, 2000] - I. Kakadiaris, D. Metaxas, *Model-based estimation of 3D human motion*, IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12), 1453-1459, 2000

- [Kambhamettu, 1994] - C. Kambhamettu, D. Goldgof, D. Terzopoulos, T. Hung, *Non-rigid Motion Analysis*, Handbook of PRIP: Computer vision, Tzay Young, San Diego, Academic Press. 2, 405/430, 1994
- [Kapur, 1996] - K. Kapur, E. Grimson, W. Wells, R. Kikinis, *Segmentation of Brain Tissue from Magnetic Resonance Images*, Medical Image Analysis 1(2), 109/127, 1996
- [Kass, 1988] - M. Kass, A. Witkin, D. Terzopoulos, *Snake: Active Contour Model*, International Journal of Computer Vision 1(4), 321/331, 1988
- [Kelly, 1995] - P. Kelly, A. Katkere, D. Kuramura, S. Moezzi, S. Chatterjee, R. Jain, *An Architecture for Multiple Perspective Interactive Video*, ACM Conference on Multimedia, 1995
- [Kervrann, 1998] - C. Kervrann, *Learning Probabilistic Deformation Models from Image Sequences*, Signal Processing 71, 155/171, 1998
- [Keserci, 2002] - B. Keserci, H. Yoshida, *Computerized Detection of Pulmonary Nodules in Chest Radiographs based on Morphological Features and Wavelet Snake Model*, Medical Image Analysis 6(4), 431/447, 2002
- [Kirubarajan, 2004] - T. Kirubarajan, Y. Bar-Shalom, *Probabilistic Data Association Techniques for Target Tracking in Clutter*, 92(3), 536/557, 2004
- [Kita, 1996] - Y. Kita, *Elastic-Model Driven Analysis of Several Views of a Deformable Cylindrical Object*, IEEE Transactions On Pattern Analysis And Machine Intelligence 18(12), 1996
- [Kojima, 2002] - A. Kojima, T. Tamura, K. Fukunaga, *Natural Language Description of Human Activities from Video Images Based on Concept Hierarchy of Actions*, International Journal of Computer Vision 50(2), 171/184, 2002
- [Kondo, 2004] - T. Kondo, S. Ong, K. Foong, *Tooth Segmentation of Dental Study Models using Range Images*, IEEE Transactions on Medical Imaging 23(3), 350/362, 2004
- [L Wang, 2003] - L. Wang, T. Tan, H. Ning, W. Hu, *Silhouette Analysis-Based Gait Recognition for Human Identification*, IEEE Transactions on Pattern Analysis and Machine Intelligence 25(12), 2003
- [Lakany, 1999] - H. Lakany, G. Hayes, M. Hazlewood, S. Hillman, *Human Walking: Tracking and Analysis*, IEEE Electronic & Communications, Colloquium on Motion Analysis and Tracking, 1999
- [Leymarie, 1993] - F. Leymarie, M. Levine, *Tracking Deformable Objects in the Plane using an Active Contour Model*, IEEE Transactions on Pattern Analysis and Machine Intelligence 15(6), 635/646, 1993

- [Li, 2004] - B. Li, Q. Meng, H. Holstein, *Articulated Pose Identification with Sparse Point Features*, IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics 34(3), 2004
- [Liebowitz, 2003] - D. Liebowitz, S. Carlsson, *Uncalibrated Motion Capture Exploiting Articulated Structure Constraints*, International Journal of Computer Vision 51(3), 171/187, 2003
- [Lipton, 1998] - A. Lipton, H. Fujiyoshi, R. Patil, *Moving Target Classification and Tracking from Real-time Video*, Proceedings 1998 DARPA Image Understanding Workshop (IUW'98), 1998
- [Little, 1998] - J. Little, J. Boyd, *Recognizing People by Their Gait: The Shape of Motion*, Videre - Journal of Computer Vision Research 1(2), 1/32, 1998
- [Lu, 2004] - C. Lu, N. Ferrier, *Repetitive Motion Analysis: Segmentation and Event Classification*, IEEE Transactions on Pattern Analysis and Machine Intelligence 26(2), 2004
- [Ma, 2004] - L. Ma, K. Khorasani, *Facial Expression Recognition Using Constructive Feedforward Neural Networks*, IEEE Transactions on Systems, Man and Cybernetics, Part B 34(3), 1588/1595, 2004
- [Marques, 2003] - J. Marques, P. Jorge, A. Abrantes, J. Lemos, *Tracking Groups of Pedestrians in Video Sequences*, IEEE Workshop on Multi-Object Tracking, Madison, 2003
- [Marroquin, 2003] - J. Marroquin, E. Santana, S. Botello, *Hidden Markov Measure Field Models for Image Segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence 25(11), 2003
- [Masoud, 2003] - O. Masoud, N. Papanikolopoulos, *Recognizing Human Activities*, Proceedings IEEE International Conference on Advanced Video and Signal Based Surveillance AVSS2003, Miami, 2003
- [Matthews, 2004] - I. Matthews, S. Baker, *Active Appearance Models Revisited*, International Journal of Computer Vision 60(2), 135/164, 2004
- [Maurincomme, 1993] - E. Maurincomme, D. Friboulet, G. Finet, I. Magnin, J. Reiber, *ADDER: A Snake-Based Segmentation Approach For Intravascular Ultrasound Images*, DICTA-93, 1993
- [McInerney, 1996] - T. Mcinerney, D. Terzopoulos, *Deformable Models in Medical Image Analysis: A Survey*, Medical Image Analysis 1(2), 1996
- [McInerney, 2002] - T. Mcinerney, G. Hamarneh, M. Shenton, D. Terzopoulos, *Deformable Organisms for Automatic Medical Image Analysis*, Medical Image Analysis 6(3), 251/266, 2002

- [Metaxas, 1993] - D. Metaxas, D. Terzopoulos, *Shape and Nonrigid Motion Estimation through Physics-Based Synthesis*, IEEE Transactions Pattern Analysis and Machine Intelligence 15(6), 580/591, 1993
- [Meyer, 1997] - D. Meyer, J. Denzler, H. Niemann, *Model Based Extraction of Articulated Objects in Image Sequences*, 4th. International Conference on Image Processing, 1997
- [Mikic, 2003] - I. Mikic, M. Trivedi, E. Hunter, P. Cosman, *Human Body Model Acquisition and Tracking using Voxel Data*, International Journal of Computer Vision 53(3), 199/223, 2003
- [Mittal, 2003] - A. Mittal, L. Davis, *M2 Tracker: A Multi-View Approach to Segmenting and Tracking People in a Cluttered Scene*, International Journal of Computer Vision 51(3), 189/203, 2003
- [Moeslund, 2001] - T. Moeslund, E. Granum, *A Survey of Computer Vision-Based Human Motion Capture*, Computer Vision and Image Understanding, 81, 231/268, 2001
- [Nascimento, 2003a] - J. Nascimento, A. Abrantes, J. Marques, *Using Middle Level Features for Robust Shape Tracking*, Pattern Recognition Letters 24(1), 295/307, 2003a
- [Nascimento, 2003b] - J. Nascimento, J. Marques, *An adaptive Potential for Robust Shape Estimation*, Image and Vision Computing Journal 21, 1107/1116, 2003b
- [Nastar, 1994] - C. Nastar, *Modèles Physiques Déformables et Modes Vibratoires pour l'Analyse du Mouvement non-rigide dans les Images Multidimensionnelles*, L'École Nationale des Ponts et Chaussées, 1994
- [Nastar, 1996] - C. Nastar, N. Ayache, *Frequency-Based Nonrigid Motion Analysis: Application to Four Dimensional Medical Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 1067/1079, 1996
- [Niyogi, 1994] - S. Niyogi, E. Adelson, *Analyzing and Recognizing Walking Figures in XYT*, Conference on Computer Vision and Pattern Recognition, 1994
- [Njastad, 1999] - J. Njastad, S. Grinaker, G. Storhaug, *Estimating Parameters in a 21/2D Human Model*, 11th. Scandinavian Conference on Image Analysis, Kangerlussuaq, Greenland, 1999
- [Olabarriaga, 2001] - S. Olabarriaga, A. Smeulders, *Interaction in the Segmentation of Medical Images: A Survey*, IEEE Transactions on Medical Imaging 5(2), 127/142, 2001
- [O'Rourke, 1980] - J. O'Rourke, N. Badler, *Model-based image analysis of human motion using constraint propagation*, IEEE Transactions on Pattern Analysis and Machine Intelligence 2(6), 522/536, 1980
- [Ozer, 2002] - L. Ozer, W. Wolf, *Real-time posture and activity recognition*, Proceedings Workshop on Motion and Video Computing, 133/138, 2002

- [Paling, 2004] - S. Paling, E. Williams, R. Barber, E. Burton, W. Crum, N. Fox, J. O'brien, *The Application of Serial MRI Analysis Techniques to the Study of Cerebral Atrophy in Late-Onset Dementia*, *Medical Image Analysis* 8(1), 69/79, 2004
- [Park, 1996] - J. Park, D. Metaxas, L. Axel, *Analysis of Left Ventricular Wall Motion based on Volumetric Deformable Models And MRI-Spamm*, *Medical Image Analysis* 1(1), 53/71, 1996
- [Park, 2004] - W. Park, D. Chaffin, B. Martin, *Toward Memory-Based Human Motion Simulation: Development and Validation of a Motion Modification Algorithm*, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 34(3), 2004
- [Pavlovic, 1997] - V.I. Pavlovic, R. Sharma, T.S. Huang, *Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7), 677/695, 1997
- [Pentland, 1991] - A. Pentland, B. Horowitz, *Recovery of nonrigid motion and structure*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(7), 715/729, 1991
- [Pers, 2002] - J. Pers, M. Bon, S. Kovacic, M. Cibila, B. Dezman, *Observation and Analysis of Large-scale Human Motion*, *Human Movement Science* 21(2), 295/311, 2002
- [Pinho, 2004] - R. Pinho, J. Tavares, *Dynamic Pedobarography Transitional Objects by Lagrange's Equation with FEM, Modal Matching and Optimization Techniques*, *ICIAR 2004 - International Conference on Image Analysis and Recognition*, Porto, Portugal, 2004
- [Plänkers, 1999] - R. Plänkers, P. Fua, N. D'apuzzo, *Automated Body Modelling from Video Sequences*, *IEEE MPEOPLE'99 - International Workshop on Modelling People*, Corfu, Greece, 1999
- [Plänkers, 2001] - R. Plänkers, P. Fua, *Articulated Soft Objects for Video-based Body Modelling*, *International Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, 2001
- [Polana, 1994] - R. Polana, R. Nelson, *Low level Recognition of Human Motion*, *Workshop on Motion of Non-Rigid and Articulated Objects*, Austin, USA, 1994
- [R Gonzalez, 2003] - R. Gonzalez, R. Woods, *Digital Image Processing*, 2003
- [Rahman, 2002] - M. Rahman, K. Nakamura, S. Ishikawa, *Recognizing human behavior using universal eigenspace*, *Proceedings. 16th International Conference on Pattern Recognition* 1, 295/298, 2002
- [Ramanan, 2003] - D. Ramanan, D. Forsyth, *Finding and Tracking People from the Bottom Up*, *IEEE Computer Vision and Pattern Recognition II*, 467/474, 2003

- [Rapantzikos, 2003] - K. Rapantzikos, M. Zervakis, K. Balas, *Detection and Segmentation of Drusen Deposits on Human Retina: Potential in the Diagnosis of Age-Related Macular Degeneration*, *Medical Image Analysis* 7(1), 95/108, 2003
- [Reynolds, 2002] - D. Reynolds, J. Riley, *Remote-sensing, Telemetric and Computer-based Technologies for Investigating Insect Movement: A Survey of Existing and Potential Techniques*, *Computers and Electronics in Agriculture* 35, 271/307, 2002
- [Rigoll, 2000] - G. Rigoll, S. Eickeler, S. Muller, *Person Tracking in Real world Scenarios using Statistical Methods*, 4th. International Conference on Automatic Face and Gesture Recognition,, Grenoble, France, 2000
- [Rohr, 1997] - K. Rohr, *Chapter 8, Human Movement Analysis Based on Explicit Motion Models*, Dordrecht Boston, Kluwer Academic Publishers, 171/198, 1997
- [Rosales, 1998] - R. Rosales, S. Sclaroff, *Improved Tracking of Multiple Humans with Trajectory Prediction and Occlusion Modelling*, *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Workshop on the Interpretation of Visual Motion*, Santa Barbara, 1998
- [Rossi, 1994] - M. Rossi, A. Bozzoli, *Tracking and Counting Moving People*, Trento, Italy, IRST, 1994
- [Sebastian, 2003] - T. Sebastian, H. Tek, J. Crisco, B. Kimia, *Segmentation of Carpal Bones from CT Images using Skeletally Coupled Deformable Models*, *Medical Image Analysis* 7(1), 21/45, 2003
- [Shechter, 2004] - G. Shechter, C. Ozturk, J. Resar, E. Mcveigh, *Respiratory Motion of the Heart from Free Breathing Coronary Angiograms*, *IEEE Transactions on Medical Imaging* 23(8), 1046/1056, 2004
- [Sidenbladh, 2000] - H. Sidenbladh, F. Torre, M. Black, *A framework for modelling the appearance of 3D Articulated Figures*, 4th International Conference on Automatic Face and Gesture Recognition, Grenoble, France, 2000
- [Sidenbladh, 2003] - H. Sidenbladh, M. Black, *Learning the Statistics of People in Images and Video*, *International Journal of Computer Vision* 54, 183/209, 2003
- [Silaghi, 1998] - M. Silaghi, R. Plankers, R. Boulic, P. Fua, D. Thalmann, *Local and Global Skeleton Fitting Techniques for Optical Motion Capture*, *International Workshop on Modelling and Motion Capture Techniques for Virtual Environments*, Geneva, Switzerland, 1998
- [Spelke, 1994] - E. Spelke, P. Vishton, C. Hofsten, *Object Perception, Object-Directed Action, and Physical Knowledge in infancy*, 165/179, 1994

- [Stauffer, 1999] - C. Stauffer, W. Grimson, *Adaptive Background Mixture Models for Real-time Tracking*, Computer Vision and Pattern Recognition, 2246/2252, 1999
- [Tavares, 1995] - J. Tavares, Dissertação de Mestrado, *Obtenção de Estrutura Tridimensional a partir de Movimento de Câmara*, Faculdade de Engenharia, Universidade do Porto, 1995
- [Tavares, 2000] - J. Tavares, Tese de Doutoramento, *Análise de Movimento de Corpos Deformáveis usando Visão Computacional*, Faculdade de Engenharia, Universidade do Porto, 2000
- [Thalmann, 1996] - D. Thalmann, J. Shen, E. Chauvineau, *Fast Realistic Human Body Deformations for Animation and VR Applications*, Computer Graphics International, Pohang, Korea, 1996
- [Tominaga, 2002] - M. Tominaga, H. Hongo, H. Koshimizu, Y. Niwa, K. Yamamoto, *Estimation of human motion from multiple cameras for gesture recognition*, Proceedings 16th International Conference on Pattern Recognition, 2002
- [Trattner, 2004] - S. Trattner, H. Greenspan, G. Tepper, S. Abboud, *Automatic Identification of Bacterial Types using Statistical Imaging Methods*, IEEE Transactions on Medical Imaging 23(7), 807/820, 2004
- [Watta, 2000] - P. Watta, N. Gandhi, S. Lakshmanan, *An eigenface approach for estimating driver pose*, Proceedings IEEE Intelligent Transportation Systems, 2000
- [Wren, 1997] - C. Wren, A. Azarbayejani, T. Darrell, A. Pentland, *Pfinder: Real-Time Tracking of the Human Body*, IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7), 1997
- [X Wang, 2004] - X. Wang, L. He, W. Wee, *Deformable Contour Method: A Constrained Optimization Approach*, International Journal of Computer Vision 59(1), 87/108, 2004
- [Xie, 2004] - X. Xie, M. Mirmehdi, *RAGS: Region-Aided Geometric Snake*, IEEE Transactions on Image Processing 13(5), 2004
- [Yamaguchi, 2003] - N. Yamaguchi, H. Mizoguchi, *Robot vision to recognize both face and object for human-robot ball playing*, IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2003
- [Yamamoto, 1991] - M. Yamamoto, K. Koshikawa, *Human Motion Analysis Based on a Robot Arm Model*, Conference on Computer Vision and Pattern Recognition, 1991
- [Yin, 2001] - X. Yin, M. Xie, *Hand Gesture Segmentation, Recognition and Application*, Proceedings 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2001
- [Zaritsky, 2003] - R. Zaritsky, N. Peterfreund, N. Shimkin, *Velocity-Guided Tracking of Deformable Contours in Three Dimensional Space*, International Journal of Computer Vision 51(3), 219/238, 2003

[Zheng, 1998] - J. Zheng, S. Suezaki, *A Model Based Approach in Extracting and Generating Human Motion*, International Conference in Pattern Recognition, 1998