



Universidade do Porto

FEUP Faculdade de
Engenharia

Relatório Interno

Breve Introdução aos Modelos Pontuais de Distribuição em Visão por Computador

Maria João Vasconcelos

Aluna de Mestrado em Estatística Aplicada e Modelação
Universidade do Porto, Faculdade de Engenharia
email: meridjo@netcabo.pt

João Manuel R. S. Tavares

INEGI – Instituto de Engenharia e Gestão Industrial, Laboratório de Óptica e Mecânica Experimental
Universidade do Porto, Faculdade de Engenharia, Departamento de Engenharia Mecânica e Gestão Industrial
Rua Dr. Roberto Frias, s/n 4200-465 PORTO – PORTUGAL
Tel. +351225081487
email: tavares@fe.up.pt
url: <http://www.fe.up.pt/~tavares>

LOME

Laboratory of Optics and Experimental Mechanics

INEGI / FEUP

Julho 2003

Resumo

O objectivo principal deste relatório é ajudar o leitor a familiarizar-se com os modelos pontuais de distribuição, a perceber a sua construção e a identificar possíveis aplicações. Assim neste relatório explica-se o processo de construção de um modelo pontual de distribuição, descreve-se a evolução que estes modelos têm sofrido desde a sua criação e apresentam-se alguns exemplos de aplicações.

O modelo pontual de distribuição (*PDM - Point Distribution Models*) dá-nos a informação da forma média de uma estrutura assim como os seus desvios admissíveis. O modelo é construído a partir de um conjunto de imagens da estrutura que queremos analisar e cada um desses objectos é representado como um conjunto de pontos-chave, pontos notáveis do contorno (ou da superfície, no caso de estruturas 3D) da estrutura. O modelo pontual de distribuição é obtido através do estudo das características do conjunto de pontos-chave: depois se emparelhar os pontos-chave de todas as estruturas do conjunto de treino é feita uma análise em componentes principais. Assim consegue-se obter um modelo pontual de distribuição com um reduzido número de parâmetros linearmente independentes que traduzem a forma média da estrutura em estudo e os desvios permitidos.

Tal como se estuda a forma da estrutura, é possível estudar os níveis de cinzento que esta estrutura pode adquirir, também através dos modelos pontuais de distribuição, e ao anexar estas duas características, obtemos um modelo combinado mais realístico para a estrutura analisada. Com as modelações anteriores podemos então construir os modelos de forma activa (*ASM - Active Shape Models*) e os modelos de aparência activa (*AAM - Active Appearance Models*), modelos que permitem, entre outras operações, identificar os objectos estudados em novas imagens.

Estes modelos estatísticos são muito úteis na interpretação de imagens, por exemplo podem ser utilizados em áreas muito distintas como: medicina, na localização de ossos em imagens médicas; indústria, na inspecção de produtos; e na segurança, no reconhecimento de faces.

Índice

1 - Construção do Modelo Pontual de Distribuição	1
2 - Evolução dos Modelos Pontuais de Distribuição.....	5
3 - Exemplos de Aplicações	7
Bibliografia.....	13

1 - Construção do Modelo Pontual de Distribuição

Com a utilização do modelo pontual de distribuição (*PDM*) pretende-se obter um modelo que represente a forma média de um dado objecto e os desvios admissíveis para a sua forma a partir de um conjunto de objectos exemplo (conjunto treino).

Assim o processo de modelação organiza-se da seguinte forma:

- 1- Etiquetagem do conjunto de treino;
- 2- Alinhamento do conjunto de treino;
- 3- Estudo das variações admissíveis;
- 4- Escolha do número de modos de variação para representar o objecto em causa.

Em seguida explica-se resumidamente em que consiste cada uma destas etapas necessárias para a construção do modelo pontual de distribuição:

1- Etiquetagem do conjunto de treino

Os objectos (2D) são representados por um conjunto de pontos notáveis etiquetados. Esses pontos devem ser colocados de forma consistente em todos os objectos do conjunto de treino; isto é, os pontos-chave devem ser colocados sensivelmente na mesma posição em cada objecto. O método mais simples para etiquetar os objectos é fazê-lo manualmente, o que implica que o utilizador que faz essa tarefa conheça bem a aplicação em estudo, de forma a escolher a localização mais apropriada de cada ponto-chave, e que seja capaz de colocá-los correctamente nas diferentes imagens de treino. Na verdade esta é uma operação muito morosa e têm sido estudados métodos automáticos ou semi-automáticos para a realização desta tarefa.

Em geral os pontos-chave são colocados nos locais que melhor descrevem objecto, normalmente nos seus limites ou outros locais representativos (pontos de valor elevado de curvatura, vértices, pontos de bifurcação, etc.).

2- Alinhamento do conjunto de treino

O método de modelação em análise consiste em estudar a variação das coordenadas dos pontos-chave ao longo do conjunto de treino. Para ser possível comparar pontos equivalentes em diferentes objectos estes devem estar alinhados.

Assim o primeiro passo será estabelecer as correspondências entre os objectos, ou seja, construir um vector de coordenadas para cada objecto da forma:

$$x_i = (x_{i0}, y_{i0}, \dots, x_{ik}, y_{ik}, \dots, x_{in-1}, y_{in-1})^T,$$

onde $i = 1 \dots N_s$ com N_s a representar o número de objectos e n o número de pontos-chave.

Como foi referido anteriormente, este passo muitas vezes é realizado manualmente mas já existem algoritmos que determinam emparelhamentos entre dois conjuntos de pontos que serão também estudados durante a dissertação.

O segundo passo será alinhar os objectos através da sua rotação, translação e redimensionamento (escalamento). O objectivo é minimizar a soma quadrática das distâncias entre pontos equivalentes.

Um possível método iterativo para este passo é o seguinte:

1. Calcular o centro de gravidade de cada exemplo e colocar todos os objectos com o mesmo centro de gravidade (origem) através de um movimento de translação.
2. Escolher um objecto como estimativa inicial da forma média e redimensioná-lo de forma a $|\bar{x}| = 1$.
3. Guardar a estimativa inicial da forma média como \bar{x}_0 .
4. Alinhar todos os objectos com a forma média.

(Duas formas dizem-se alinhadas quando D é mínimo:

$$D = \sum_{i=1}^{N_s} |x_i - \bar{x}|^2 .)$$

5. Re-estimar a forma média a partir das formas anteriores.
6. Alinhar a nova forma média, \bar{x} , com \bar{x}_0 e redimensioná-la de forma que $|\bar{x}| = 1$.
7. Repetir os passos anteriores até que o valor da forma média convirja.

3- Estudo das variações admissíveis

Suponha-se agora que se tem N_s conjuntos de pontos x_i alinhados. A forma média calcula-se por:

$$\bar{x} = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i .$$

Os modos de variação, isto é, as formas segundo as quais os pontos tendem a mover-se, podem ser encontradas aplicando uma análise em componentes principais aos desvios da média.

A referida análise em componentes principais consiste em calcular os vectores e valores próprios da matriz de covariâncias S definida por:

$$S = \frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})(x_i - \bar{x})^T .$$

Os modos de variação dos pontos-chave são descritos pelos vectores próprios de S , p_i , tal que:

$$Sp_i = \lambda_i p_i ,$$

onde λ_i é o $i^{\text{ésimo}}$ valor próprio de S ($\lambda_i \geq \lambda_{i+1}$) e $p_i^T p_i = 1$.

Sabe-se que os vectores próprios da matriz de covariâncias, S , correspondentes aos valores próprios mais elevados descrevem a maior parte das variações admissíveis, e a proporção da variância total explicada por cada vector próprio é igual ao valor próprio correspondente. Assim a maior parte da variância de forma pode ser explicada por um pequeno número de vectores próprios t , chamados modos de variação.

Desta forma, cada objecto do conjunto de treino pode ser descrito pela forma média e pela combinação dos primeiros t vectores próprios obtidos:

$$x = \bar{x} + Pb ,$$

onde $P = (p_1 p_2 \dots p_t)$ é a matriz dos primeiros t vectores próprios e $b = (b_1 b_2 \dots b_t)$ é o vector de pesos de cada vector próprio.

Os vectores próprios são ortogonais e portanto $P^T P = I$ e:

$$b = P^T (x - \bar{x}) .$$

As equações anteriores permitem criar novos objectos variando os parâmetros b_i dentro de limites aceitáveis.

Como a variância de b_i sobre o conjunto de treino é dada por λ_i , os seus limites são da ordem:

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i},$$

pois a população está concentrada entre três desvios padrões da média.

4- Escolha do número de modos de variação

Esta etapa corresponde ao número de modos de variação, t , que se quer reter, este pode ser escolhido de diversas maneiras, de acordo com a aplicação em que o modelo será utilizado.

A maneira mais simples é escolher a percentagem de variância explicada pelo modelo. Sabe-se que a soma de todos os valores próprios, que representam o conjunto de treino, é igual à variância total dos dados:

$$V_t = \sum \lambda_i.$$

Assim pode-se calcular t a partir de:

$$\sum_{i=1}^t \lambda_i = f_V V_t,$$

onde f_V define a proporção do total de variância que se quer explicar.

Consoante a aplicação o número de modos, t , é escolhido; geralmente, escolhe-se t de forma a que a variância explicada seja elevada (> 90%).

Uma alternativa é calcular o valor t mínimo que representa os dados satisfatoriamente. Para tal, constrói-se os modelos aumentando o número de modos de variação, t , até que o valor t corresponda ao critério seleccionado.

Desta forma consegue-se então determinar as variações principais de uma classe de formas partindo da representação de objectos como um conjunto de pontos etiquetados, ou seja, deste modo calcula-se o modelo pontual de distribuição.

2 - Evolução dos Modelos Pontuais de Distribuição

Após a criação dos modelos pontuais por (Cootes, Taylor et al. 1992), estes têm vindo a ser aperfeiçoados ao longo dos anos por diferentes grupos de investigadores. Em seguida apresenta-se sucintamente as evoluções mais importantes deste modelo.

Em (Cootes and Taylor 1993) é descrito um método para localizar estruturas em imagens (*ASM - Active Shape Models*) através do estudo dos modelos, os chamados modelos pontuais de distribuição, criados a partir da localização dos pontos notáveis do contorno da estrutura e dos níveis de cinzento que estes pontos apresentam. A adição dos níveis de cinzento dos objectos em estudo permitiu assim obter melhores modelos para essas estruturas.

O processo para obter os *Active Shape Models* consiste no seguinte: depois de gerar um modelo da forma e de ter uma descrição sobre os níveis de cinzento de cada modelo pontual tem-se como objectivo localizar exemplos da estrutura modelada em novas imagens. O processo considerado é constituído por duas etapas:

- São calculadas várias hipóteses sobre as possíveis localizações dos pontos notáveis;
- Cada hipótese é estudada e refinada sendo escolhida a melhor delas.

Para ultrapassar o problema da colocação manual dos pontos-chave e a determinação da correspondência pontual, (Hill and Taylor 1994) apresenta um método para a determinação automática dos pontos (2D) a serem utilizados na geração dos modelos pontuais de distribuição e um método para a determinação da correspondência pontual, sendo apresentados resultados em imagens do coração e da mão.

Em (Taylor, Cootes et al. 1995) é descrita uma técnica para construir modelos compactos da forma e aparência de objectos deformáveis (por exemplo órgãos) presentes em imagens a duas e a três dimensões. É descrito também como gerar modelos 3D e como estes são usados para a segmentação de imagens do cérebro a 3D de ressonância magnética.

Mais tarde, em (Cootes, Edwards et al. 1998) são apresentados os *Active Appearance Models* (*AAM*). Um *AAM* contém um modelo estatístico da forma e aparência do objecto em estudo, podendo este modelo ser generalizado para qualquer exemplo válido. Durante uma fase de

treino estuda-se a relação entre os parâmetros obtidos pelo modelo e os erros residuais existentes entre a imagem treino e o modelo produzido com o objectivo de produzir um modelo final mais robusto.

A comparação dos *ASM's* com os *AAM's* feita em (Cootes, Edwards et al. 1999) sugere que se deve estudar um novo modelo que contenha uma combinação dos dois modelos anteriores de forma a juntar as vantagens de cada um deles. Se por um lado os *ASM* abrangem uma maior área de procura do que os *AAM*, estes conseguem construir um modelo com um reduzido número de pontos-chave e como a etiquetagem das formas é um processo complexo, quanto menor for o número de pontos notáveis necessários melhor.

Em (Cootes, Walker et al. 2000) demonstra-se que um pequeno número de modelos estatísticos a 2D são suficientes para capturar a forma e aparência de uma face a partir de qualquer ponto de vista (frontal e lateral). Os modelos são lineares e são capazes de localizar faces em novas imagens utilizando o algoritmo *AAM*.

Os modelos estatísticos (*statistical shape models*) têm assim sido utilizados como base para a segmentação e interpretação de imagens em diversas aplicações. Ultimamente tem-se usado *ASM* robustos ((Rogers and Graham 2002) e (Rogers and Graham 2002)) para que as aplicações, principalmente na área da análise de imagens médicas, sejam mais eficientes e ajustáveis.

3 - Exemplos de Aplicações

Em primeiro lugar mostra-se um exemplo de um modelo pontual de distribuição, este exemplo foi retirado de (Davies and Cootes 1999). A forma modelada é uma mão. Foram consideradas várias figuras de uma mão como conjunto de treino, *figura 1*, cada objecto foi representado por 72 pontos-chave.

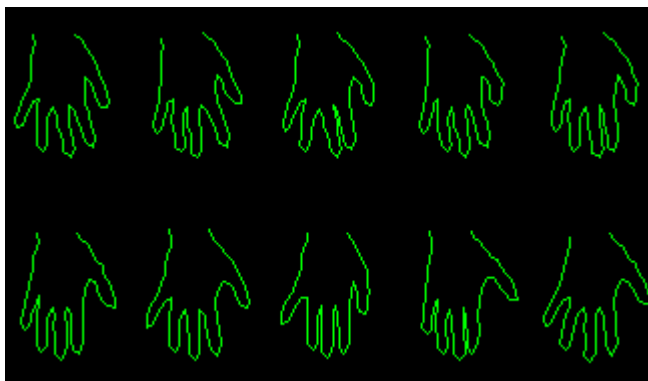


Fig 1 – Alguns exemplos do conjunto de treino.

Depois de se alinhar o conjunto de treino, estuda-se as variações admissíveis do modelo através de uma análise em componentes principais. Os três primeiros modos de variação do modelo pontual de distribuição são apresentados na *figura 2*.

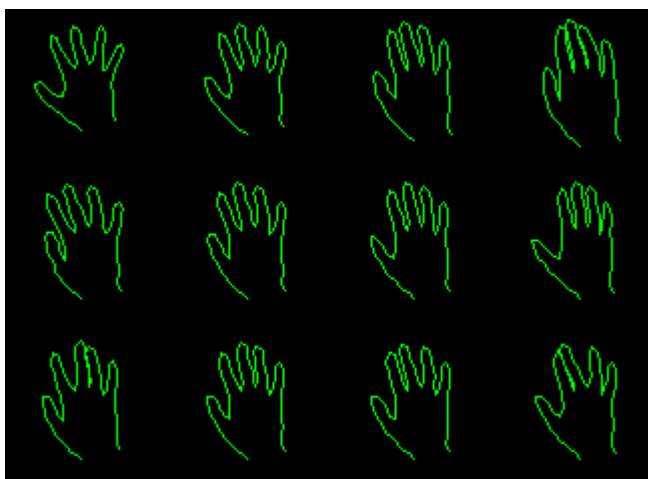


Fig 2 – Primeiros três modos de variação do modelo pontual.

A primeira aplicação que se apresenta é a identificação de uma estrutura facial, numa imagem através da utilização do processo iterativo dos *Active Shape Models*. Cada estrutura facial foi representada por 68 pontos-chave, *figura 3*, e foi construído um modelo a partir de um conjunto de várias imagens. O primeiro modo de variação do modelo é apresentado na *figura 4* e corresponde à rotação da cabeça.

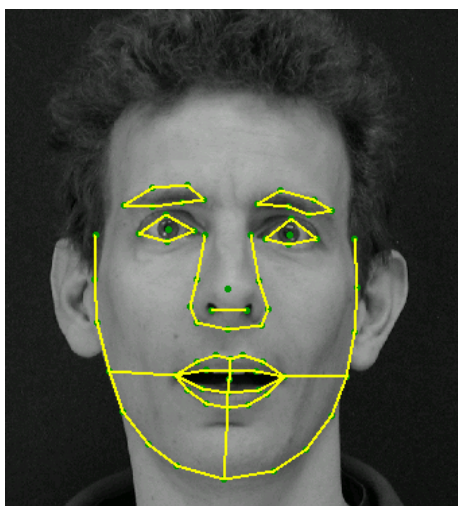


Fig 3 – Estrutura da face representada por 72 pontos-chave.

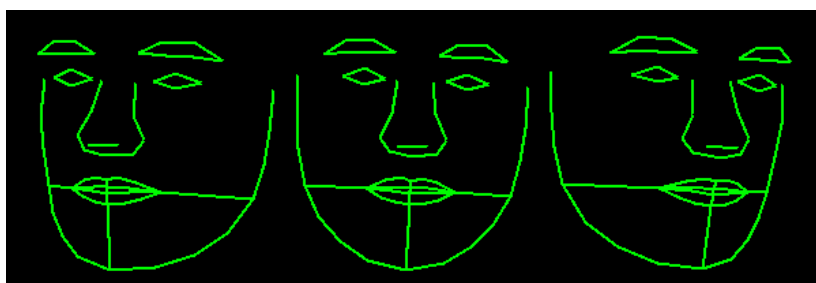


Fig 4 – Primeiro modo de variação.

Após ter-se construído o modelo, foi utilizado o *ASM* para encontrar a estrutura facial numa nova imagem. Na *figura 5* pode-se ver como o processo converge, identificando correctamente a face da imagem em estudo.



Fig 5 – Posição inicial.



Fig 6 – Após 5 iterações.



Fig 7 – Convergência.

Em seguida apresenta-se um modelo de uma face construído com algoritmo do *Active Appearance Model*, este exemplo foi retirado de (Cootes, Edwards et al. 1998). Para construir o modelo foram utilizadas 400 imagens de faces, e cada uma foi etiquetada com 122 pontos-chave, *Figura 8*. O modelo da forma gerado tem 23 parâmetros, o modelo dos níveis de cinzento (*shape-free grey model*) tem 114 parâmetros e o modelo de aparência gerado necessita apenas de 80 parâmetros para explicar 98% da variância total observada.

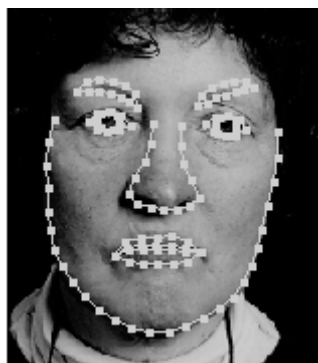


Fig 8 – Exemplo da imagem de uma face etiquetada com 122 pontos-chave.

As *figuras 9 e 10* mostram o efeito da variação dos dois primeiros parâmetros do modelo gerado entre ± 3 desvios padrões da média determinada inicialmente. A *figura 11* mostra o efeito da variação dos primeiros quatro parâmetros do modelo de aparência, estes parâmetros revelam mudanças de identidade, pose e expressão.



Fig 9 – Primeiros dois modos de variação da forma ($\pm 3sd$).



Fig 10 – Primeiros dois modos de variação dos níveis de cinzento ($\pm 3sd$).



Fig 11 – Primeiros quatro modos de variação da aparência ($\pm 3sd$).

Na medicina também é possível aplicar os *PDM's* com sucesso, o exemplo seguinte aplica os *Active Appearance Models* para a localização de um osso (o metacarpo) em imagens médicas. A *figura 12* mostra o conjunto de pontos-chave necessário para a construção do modelo final.

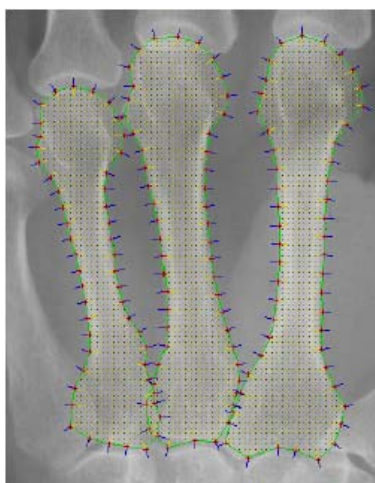


Fig 12 – Exemplo de uma imagem etiquetada.

As figuras 13 e 14 mostram os primeiros modos de variação dos modelos calculados da forma e textura.

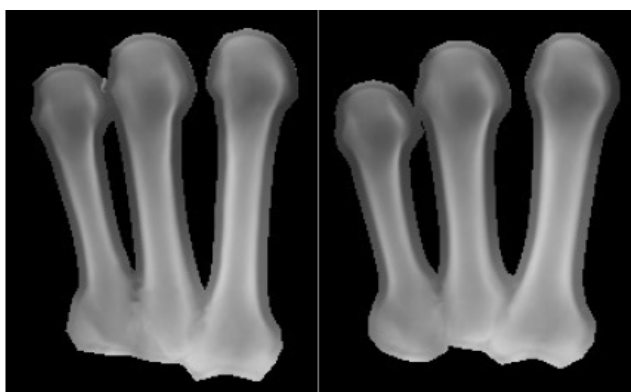


Fig 13 – Primeiro modo de variação da forma ($\pm 3sd$).

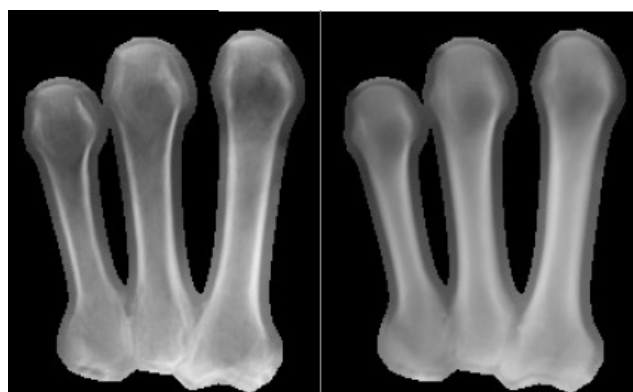


Fig 14 – Primeiro modo de variação da textura ($\pm 3sd$).

A figura 15 mostra como se pode utilizar os *AAM's* na identificação de estruturas em imagens médicas. Por vezes esta é uma tarefa difícil pois muitas vezes as imagens têm pouca qualidade, o que pode impedir o médico de manualmente realizar uma boa avaliação.

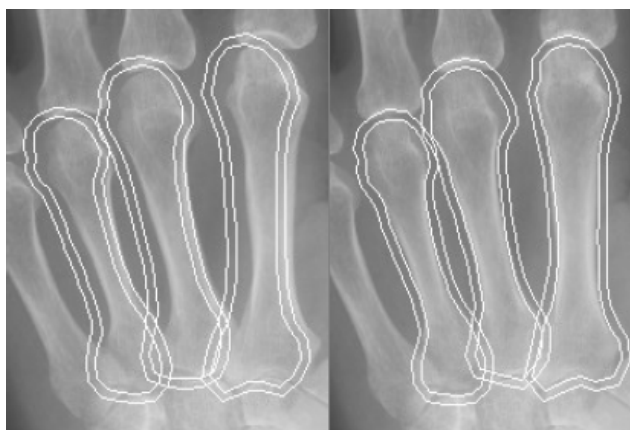


Fig 15 – Procura *AAM*, inicial e óptima.

Um outro exemplo da aplicação dos *PDM's* na área da medicina é o apresentado em seguida (Cootes, Edwards et al. 1999). Aqui utilizam-se os *Active Shape Models* para construir um modelo da estrutura central do cérebro (ventrículos, núcleos caudados e núcleos lentiformes) em imagens obtidas por ressonância magnética, na figura 16 mostra-se uma destas imagens etiquetadas com 123 pontos-chave.

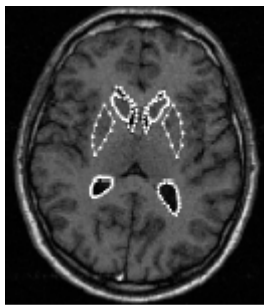


Fig 16 – Exemplo de uma imagem etiquetada.

A *figura 17* mostra como o processo dos *Active Shape Models* converge de forma a identificar a estrutura estudada numa nova imagem.

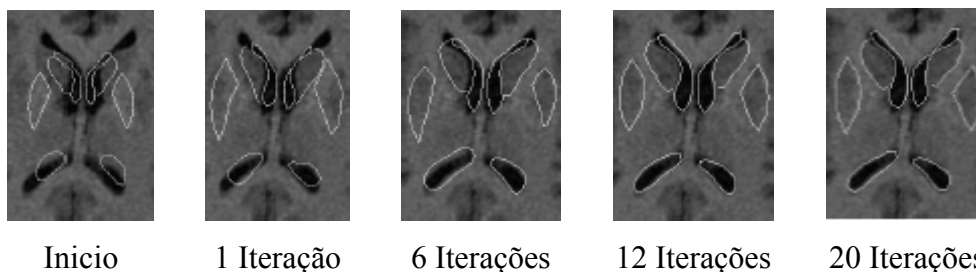


Fig 17 – Procura numa nova imagem de uma dada estrutura utilizando o *ASM*.

Em resumo, os modelos pontuais de distribuição são muito úteis na interpretação de imagens. Estes modelos são usados nos *Active Shape Models* e nos *Active Appearance Models* que iterativamente os posicionam em novas imagens. Como se verificou, os *PDM's* podem ser utilizados em diversas áreas tais como:

- Medicina – localização de ossos e órgãos em imagens médicas, muitas vezes esta é uma tarefa difícil manualmente e os *PDM's* obtêm bons resultados quer em imagens 2D quer em imagens 3D;
- Industrial – inspeção de produtos industriais;
- Segurança – identificação e reconhecimento de faces;
- ...

Bibliografia

- Cootes, T. F., G. Edwards, et al. (1999). Comparing Active Shape Models with Active Appearance Models. Department of Medical Biophysics, Manchester University.
- Cootes, T. F., G. J. Edwards, et al. (1998). Active Appearance Models. Proc. European Conference on Computer Vision.
- Cootes, T. F. and C. J. Taylor (1993). Active Shape Model Search using Local Grey-Level Models: A Quantitative Evaluation. British Machine Vision Conference, BMVA Press: 639/648.
- Cootes, T. F., C. J. Taylor, et al. (1992). Training Models of Shape from Sets of Examples. Department of Medical Biophysics, University of Manchester.
- Cootes, T. F., K. Walker, et al. (2000). View-Based Active Appearance Models. Dept. Imaging Science and Biomedical Engineering, University of Manchester.
- Davies, K. and T. Cootes (1999). Statistical Shape Models, Wolfson Image Analysis Unit - University of Manchester.
- Hill, A. and C. J. Taylor (1994). Automatic Landmark Generation for Point Distribution Models. Department of Medical Biophysics, University of Manchester.
- Rogers, M. and J. Graham (2002). Robust Active Shape Model Search. Division of Imaging Science and Biomedical Engineering, University of Manchester.
- Rogers, M. and J. Graham (2002). Robust Active Shape Model Search for Medical Image Analysis. Division of Imaging Science and Biomedical Engineering, University of Manchester.
- Taylor, C. J., T. F. Cootes, et al. (1995). Medical Image Segmentation Using Active Shape Models. Department of Medical Biophysics, University of Manchester.