

## APLICAÇÃO DO ALGORITMO ACTIVE SHAPE MODELS EM IMAGENS DE BASES NASAIS PARA AUXILIAR ANÁLISES MÉDICAS

HUGO SANTOS PIAUILINO NETO\*, KELSON RÔMULO TEIXEIRA AIRES\*, RICARDO TELES FREITAS\*,  
RODRIGO DE MELO SOUZA VERAS\*, CORNÉLIA JANAYNA PEREIRA PASSARINHO\*, VICTOR EULÁLIO  
DE SOUSA CAMPELO†

\**Departamento de Computação  
Universidade Federal do Piauí  
Teresina, Piauí, Brazil*

†*Departamento de Medicina Especializada  
Universidade Federal do Piauí  
Teresina, Piauí, Brazil*

Emails: hugos94@gmail.com, kelson@ufpi.edu.br, ricard0tfreitas@gmail.com,  
rveras@ufpi.edu.br, janaynapassarinho@gmail.com, dr.vcampelo@gmail.com

**Abstract**— This paper proposes the use of the Active Shape Models algorithm (Cootes and Taylor, 1992), to locate a set of points of interest in frontal images of nasal bases. The goal is to provide useful and accurate information for aesthetic analysis. The algorithm was implemented and tested on a set of 180 images and showed an average error of 20 pixels in relation to the markings made by a specialist.

**Keywords**— Points, Nasal, Aesthetic, Analysis, Planning, Surgical.

**Resumo**— Este trabalho propõe a utilização do algoritmo Active Shape Models (Cootes and Taylor, 1992), para a localização de um conjunto de pontos de interesse em imagens frontais de bases nasais. O objetivo é fornecer informações úteis e precisas para análises estéticas. O algoritmo foi implementado e testado em um conjunto de 180 imagens e apresentou um erro médio de 20 pixels em relação as marcações realizadas por um especialista.

**Palavras-chave**— Pontos, Nasal, Estética, Análise, Planejamento, Cirúrgico.

### 1 Introdução

Imagens faciais humanas apresentam elementos complexos que variam significativamente devido a existência de uma grande diversidade de formas, expressões, poses, planos de fundo e iluminações.

Podem ser identificados, na face humana, diversos pontos de interesse que geralmente armazenam inúmeras informações e características sobre determinada face de uma pessoa.

Cirurgiões plásticos precisam estar cientes do maior número possível de parâmetros e procedimentos necessários, para corrigir ou aprimorar a face de um paciente. Assim, eles precisam obter precisamente todos os pontos de interesse, suas relações espaciais e outros possíveis dados importantes.

A localização automática desses pontos de interesse na face é um campo de pesquisa bastante ativo na área de visão computacional e apresenta diversas aplicações (Yang et al., 2002; Jafri and Arabnia, 2009).

A tarefa de identificar essas estruturas particulares geralmente é uma tarefa árdua, onde métodos podem ser frustrados pela variabilidade inerente das formas e elementos presentes. Portanto, são necessários sistemas que possam trabalhar com variabilidade para analisar a complexidade das formas contidas nas imagens.

Este artigo concentra-se na tarefa específica

de localizar pontos de interesse em imagens faciais de base nasal. Posteriormente, esses pontos serão utilizados no auxílio de análises médicas. Foi utilizado o algoritmo *Active Shape Models* (ASM) (Cootes and Taylor, 1992), com o propósito de identificar os pontos de referência definidos neste trabalho.

Este trabalho foi dividido em sete seções, incluindo esta. A seguir, a Seção 2 apresenta uma breve descrição do algoritmo ASM. A Seção 3 descreve o cenário geral dos trabalhos relacionados. Na Seção 4 são apresentadas as informações sobre os pontos de interesse utilizados. A Seção 5 descreve a metodologia aplicada para localizar os pontos de referência. Na Seção 6, os testes e resultados da aplicação do método são discutidos. Por fim, na Seção 7, a conclusão e os trabalhos futuros são apresentados.

### 2 Active Shape Models (ASM)

Um ponto de interesse é uma posição distinguível de um elemento que pode armazenar informações e definir determinadas características. Um exemplo disso é a localização da ponta do nariz que pode ser facilmente identificada em uma imagem nasal. Diversas características faciais podem ser identificadas ao localizar pontos de interesse na face humana.

Em cada imagem, pontos de interesse podem

ser agrupados para gerar uma forma. As formas de todas as imagens são alinhadas entre si aplicando transformações de similaridade (translação, escala e rotação) para minimizar a média da distância euclidiana entre os pontos das formas. Ao final, uma forma global é gerada a partir da média das formas alinhadas de cada imagem.

O algoritmo *Active Shape Models* (ASM) utiliza, em sua execução, o conceito de multi-resolução que pode ser idealizado por uma pirâmide de uma mesma imagem com diferentes resoluções. Neste trabalho, foram utilizados três níveis de resolução na pirâmide. A primeira execução ocorre em uma proporção de 25% da imagem original. As execuções posteriores ocorrem em 50% e 100% da imagem original, respectivamente.

Durante sua execução, o ASM realiza os seguintes passos, em cada nível da pirâmide, até a forma inicial da imagem convergir para a forma global do modelo:

- (i) Agrupa os pontos de interesse da imagem gerando uma forma provisória inicial;
- (ii) Deforma a forma provisória gerada até assemelhar-se com a forma global.

Os modelos de forma individuais gerados não são confiáveis. Porém, o agrupamento do conjunto de modelos de forma individuais produzem um classificador geral mais robusto. Toda a operação é repetida em cada nível da pirâmide da imagem, da granularidade grossa para a fina.

Segue-se os dois tipos de sub-modelos que compõem o ASM: o modelo de ponto e o modelo de forma.

**Modelo de ponto** (um para cada ponto de interesse em cada nível da pirâmide) é utilizado para identificar a localização aproximada de um ponto de referência por combinação de modelos.

Qualquer *template matcher* (combinador de modelos) pode ser utilizado, mas foi escolhido um vetor gradiente normalizado com tamanho fixo (denominado de perfil), através da amostragem de uma imagem ao longo de uma reta ortogonal à forma na fronteira do ponto de interesse, denominada de *whisker*.

Utilizando as marcações realizadas pelo especialista, para cada ponto de interesse é calculado o vetor do ponto médio  $\bar{\mathbf{g}}$  e a matriz de covariância do perfil  $\mathbf{S}_{\mathbf{g}}$ . Logo após, o ponto de referência é deslocado ao longo do *whisker* para o pixel cujo ponto  $\mathbf{g}$  tem a menor distância de *Mahalanobis* do ponto médio  $\bar{\mathbf{g}}$  da forma global:

$$Mahalanobis = (\mathbf{g} - \bar{\mathbf{g}})^T \mathbf{S}_{\mathbf{g}}^{-1} (\mathbf{g} - \bar{\mathbf{g}}). \quad (1)$$

A Figura 1 apresenta o modelo de ponto sendo alinhado na estrutura da imagem. A parte quadriculada representa a estrutura alvo que o método irá identificar, os círculos representam os pontos de interesse e as retas que conectam cada círculo representam a forma. O *whisker* é representado pela

reta ortogonal à forma da fronteira, onde toda a sua extensão delimita a variação que o modelo de ponto pode alcançar.

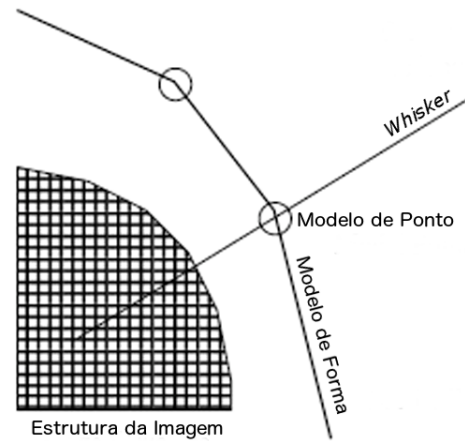


Figura 1: Modelo de ponto sendo alinhado na estrutura da imagem.

**Modelo de forma** é construído a partir de um conjunto de modelos de pontos. É gerada uma forma  $\hat{\mathbf{x}}$  com

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \Phi \mathbf{b}, \quad (2)$$

onde  $\bar{\mathbf{x}}$  é a forma média,  $\mathbf{b}$  é um vetor parâmetro e  $\Phi$  é a matriz dos autovetores selecionados da matriz de covariância  $\mathbf{S}_{\mathbf{s}}$  dos pontos das formas alinhadas. Utilizando uma abordagem de componentes principais, a variação do conjunto pode ser modelada pela ordenação dos autovalores de  $\mathbf{S}_{\mathbf{s}}$ , mantendo um número correspondente de autovetores. Um modelo de forma global único é utilizado para todo o ASM, mas deve ser redimensionado para cada nível da pirâmide.

Várias formas podem ser geradas com a Equação 2 pela variação do parâmetro  $\mathbf{b}$ . Mantendo os elementos de  $\mathbf{b}$  limitados (determinado durante a construção do modelo) garante-se que as formas geradas são realistas.

Portanto, dada uma forma sugerida  $\mathbf{x}$ , podemos calcular o parâmetro  $\mathbf{b}$  permitindo que a Equação 2 melhor aproxime  $\mathbf{x}$  com o modelo de forma  $\hat{\mathbf{x}}$ . Cootes and Taylor (2004) descrevem um algoritmo iterativo que dado  $\mathbf{b}$  e  $\mathbf{T}$  minimiza a distância total:

$$distance(\mathbf{x}, \mathbf{T}(\bar{\mathbf{x}} + \Phi \mathbf{b})), \quad (3)$$

onde  $\mathbf{T}$  é a transformação de similaridade que mapeia o espaço do modelo para o espaço da imagem.

A Figura 2 apresenta o modelo de forma ajustando-se a forma alvo desejada. Ele é inserido na imagem com a característica alvo e começa interagir com a mesma, deformando-se, até tornar seu corpo semelhante a característica desejada. Operações de translação, escala e rotação podem ser aplicadas para o modelo ajustar-se a forma completamente.

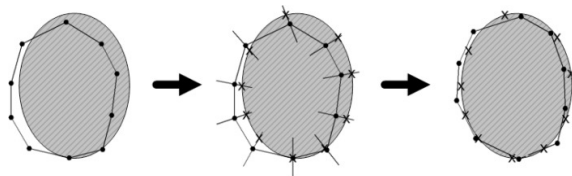


Figura 2: Modelo de forma ajustando-se a forma alvo desejada.

### 3 Trabalhos Relacionados

O algoritmo *Active Shape Models* (ASM) pertence à classe de modelos que após a forma ser inserida perto da imagem com a característica objetivo, interage com a imagem, deformando-se, até tornar sua forma semelhante à característica desejada. Eles são modelos deformáveis como *Snakes* (Kass et al., 1988), entretanto, utilizam um modelo de forma explícita para colocar restrições globais sobre a forma gerada. O ASM foi inicialmente apresentado por Cootes and Taylor (1992). Logo após, foi produzida uma sucessão cumulativa de artigos relacionados ao ASM (Cootes et al., 1995; Cootes and Taylor, 1997; Cootes et al., 1999; Cootes and Taylor, 2004).

Foram propostas muitas alterações do ASM clássico. Cootes and Taylor (1997) empregaram um modelo na forma de uma mistura de gaussiana multivariada, ao invés da distribuição gaussiana simples implícita na ASM clássica. Rogers and Graham (2002) aplicaram uma técnica de mínimos quadrados para minimizar os resíduos entre a forma do modelo e a forma sugerida, deixando o ASM mais robusto.

Zhou et al. (2003) estimam parâmetros de forma e pose utilizando inferência *Bayesiana* depois de projetar as formas em um espaço tangente. Li and Ito (2005) construíram modelos com textura utilizando o classificador de histograma *Ada-Boosted*. Já o *Active Appearance Model* (AAM) (Cootes et al., 2001), funde os modelos de forma e perfil do ASM clássico em um único modelo de aparência. Cootes et al. (1999) relatam que a precisão de localização dos pontos de interesse é melhor no geral para ASM do que para AAM, embora isso possa ter mudado com os desenvolvimentos subsequentes para o AAM.

### 4 Pontos de Interesse Nasais

Cirurgiões plásticos podem ter, segundo características próprias de cada paciente, um conjunto variado de pontos de referência necessários para realizar uma análise pré-operatória. Porém, na literatura, existem determinados pontos de interesse que frequentemente são utilizados pelos cirurgiões. Bugaighis et al. (2014) definiu diversos pontos de interesse na região nasal. Para este trabalho foram escolhidos 18 dos principais pontos

de interesse da base nasal definidos por Bugaighis et al. (2014).

A Figura 3 apresenta a localização aproximada dos pontos escolhidos em uma imagem de uma base nasal. A Tabela 1 mostra a respectiva legenda da Figura 3.

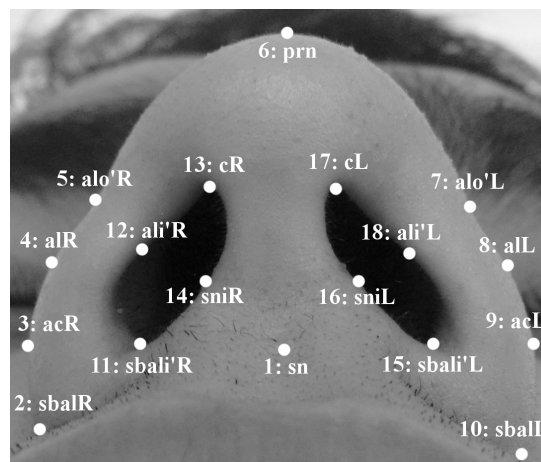


Figura 3: Localização dos 18 principais pontos de interesse da base nasal utilizados neste trabalho.

Tabela 1: Legenda dos 18 pontos de interesse apresentados na Figura 3.

| Posição      | Ponto                   |
|--------------|-------------------------|
| 1 - sn       | Subnasale               |
| 2 - sbalR    | Subalare (Right)        |
| 3 - acR      | Alare Curvature (Right) |
| 4 - alR      | Alare (Right)           |
| 5 - alo'R    | Outer Alare (Right)     |
| 6 - prn      | Pronasale               |
| 7 - alo'L    | Outer Alare (Left)      |
| 8 - alL      | Alare (Left)            |
| 9 - acL      | Alare Curvature (Left)  |
| 10 - sbalL   | Subalare (Left)         |
| 11 - sbali'R | Subalare Inner (Right)  |
| 12 - ali'R   | Inner Alare (Right)     |
| 13 - cR      | Columellar High (Right) |
| 14 - sniR    | Subnasale Inner (Right) |
| 15 - sbali'L | Subalare Inner (Left)   |
| 16 - sniL    | Subnasale Inner (Left)  |
| 17 - cL      | Columellar High (Left)  |
| 18 - ali'L   | Inner Alare (Left)      |

### 5 Metodologia

O algoritmo ASM foi minimamente modificado para atender as restrições do problema e localizar características em imagens faciais de bases nasais. Estas características podem ser distinguidas por padrões de contorno que foram identificados e utilizados para buscar os pontos de interesse necessários para a análise estética.

O método utilizado foi implementado para ser executado em imagens que obedecem determinadas condições e tenham sido adquiridas em um ambiente controlado:

- Imagem deve conter apenas uma base nasal;
- A base nasal deve estar na posição horizontal;
- Pontos da Figura 3 precisam estar visíveis;

Pode-se argumentar que essas condições são muito limitantes e não são frequentemente encontradas no mundo real. Porém, para o objetivo geral de identificar os pontos de interesse da base nasal, obedecer essas regras é de extrema importância.

Alguns fatores interferem diretamente na criação do modelo gerado pelo algoritmo. Dentre eles estão:

- **Heterogeneidade** – quanto mais heterogêneo for o conjunto de imagens utilizadas na construção do modelo, mais abrangente a identificação de pontos de interesse será;
- **Quantidade de Imagens** – uma maior quantidade de imagens gera um modelo com maior capacidade de generalização, como demonstrado na Figura 4(c), onde praticamente a sobreposição de formas abrange um grande conjunto de possibilidades para bases nasais;
- **Tamanho das Imagens** – em um cenário ideal para o ASM, imagens teriam as mesmas dimensões em largura e altura. Entretanto, quando imagens de diferentes dimensões são utilizadas na criação do modelo, poderemos ter tamanhos discrepantes entre as formas geradas. Na Figura 4 temos três modelos criados por imagens com largura de 400 *pixels* e altura variável. Apesar de existirem bases nasais de diferentes tamanhos, não pode-se afirmar que as formas mais externas das Figuras 4(b) e 4(c) são de bases nasais grandes ou de imagens com altura superior as demais;
- **Quantidade de Pontos** – o número de pontos utilizados para a criação do modelo é diretamente proporcional à acurácia da aplicação. Com uma maior quantidade de pontos, podemos representar melhor as curvas das formas. Para este trabalho, foram escolhidos 18 pontos de interesse, como apresentado na Seção 4. Porém, para a construção do modelo, foram utilizados 36 pontos para melhor representar a forma da base nasal, como demonstrado na Figura 4(a). A base de imagens utilizada continha apenas os 18 pontos iniciais identificados, por isso, foi necessário realizar a adição dos outros pontos. Este processo demandou muito tempo para ser executado pelo especialista;

- **Posicionamento dos Pontos** – uma maior exatidão durante a marcação dos pontos de interesse utilizados, gera um modelo melhor que poderá adequar-se ao posicionamento correto dos pontos durante a fase de execução da aplicação.



(a) 1 imagem.



(b) 60 imagens.



(c) 100 imagens.

Figura 4: Modelos construídos utilizando a base de imagens.

## 6 Resultados

O método foi implementado na linguagem *open source Swift* disponibilizada pela Apple (2016). Para avaliar o seu desempenho, foram gerados dois modelos de forma utilizando 60 e 100 imagens de bases nasais diferentes, respectivamente. Os testes foram realizados no conjunto total de 180 imagens, respeitando as condições impostas na Seção 5. Os

resultados obtidos foram comparados com as marcações dos pontos realizadas pelo especialista.

A distância euclidiana foi escolhida para medir o erro entre os pontos de interesse localizados manualmente e automaticamente. Antes de compilar os resultados, as distâncias devem ser colocadas em perspectiva. As reais distâncias entre as marcações podem enganar, pois apesar das imagens possuírem a mesma largura, as alturas são variáveis. Isto indica que as diferenças podem ser mais ou menos significativas dependendo do tamanho da imagem.

Foram realizadas 10 simulações para os dois modelos de forma gerados. Nos resultados apresentados neste artigo, são mostrados os intervalos de confiança considerando um nível de confiança de 95%.

Na Figura 5, o contorno mais claro (branco) representa o modelo de forma médio que foi construído utilizando 60 imagens e o mais escuro (verde) representa a melhor estimativa gerada ao fim de todas as iterações para a base nasal em questão. Podemos verificar que os pontos *Subalare* (*Right*) e (*Left*) não obtiveram uma boa aproximação dentre todos os outros pontos de interesse.

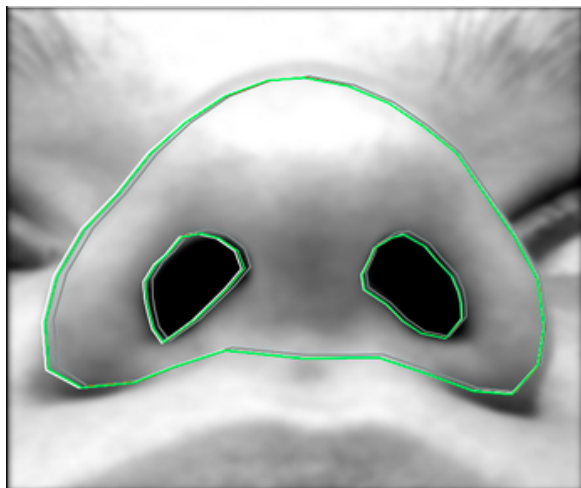


Figura 5: Teste unitário realizado com um modelo construído a partir de 60 imagens.

Na Figura 6, o contorno mais claro (branco) representa o modelo de forma médio que foi construído utilizando 100 imagens e o mais escuro (verde) é a melhor estimativa gerada após as iterações da aplicação para a imagem em questão.

A Figura 7 apresenta os gráficos contendo as probabilidades de erros médios (em *pixels*) para todos os pontos de interesse utilizados na geração dos modelos de forma de 60 e 100 imagens, respectivamente. O erro médio para a marcação de pontos por especialistas fica em torno de 4 ou 6 *pixels*. Os erros médios obtidos pelo método variam entre 14 e 22 *pixels* para cada ponto.

Observando a Figura 7, podemos perceber que o modelo gerado com 100 imagens apresenta maior

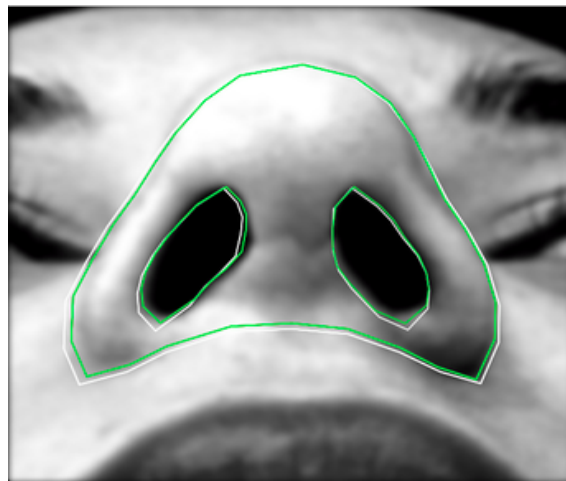
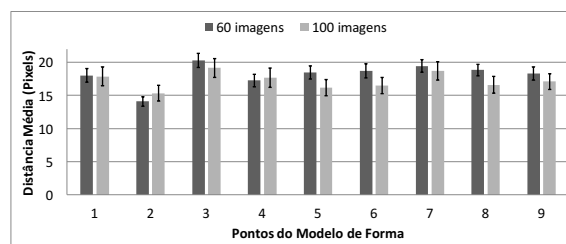


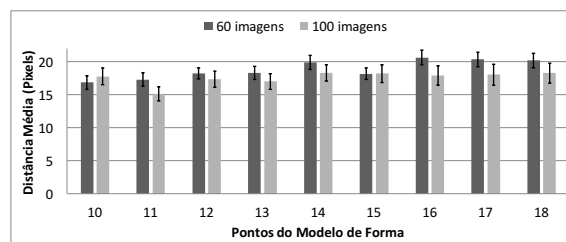
Figura 6: Teste unitário realizado com um modelo construído a partir de 100 imagens.

abrangência em comparação ao de 60 imagens. Desse modo, conforme a quantidade de imagens aumenta, o erro médio diminui.

Neste trabalho foram utilizados apenas modelos gerados a partir de pequenos conjuntos de imagens, devido a falta de bases relacionadas ao problema.



(a) Comparação dos erros médios obtidos para os pontos de 1 até 9.



(b) Comparação dos erros médios obtidos para os pontos de 10 até 18.

Figura 7: Erros médios (em *pixels*) dos modelos de 60 e 100 imagens.

Outro fator que tem influência direta sobre a taxa de erros é a quantidade de pontos de interesse utilizados na construção do modelo. Milborrow and Nicolls (2008) demonstram que para um dado número (entre 3 e 68) de pontos de referência, os que obtiveram melhores resultados foram os que continham mais pontos. Porém, o custo do tempo de procura aumenta aproximadamente de maneira linear juntamente com o número de pon-

tos de interesse. Paralelamente, o uso de memória cresce consideravelmente.

Conforme já observado na Seção 5, a quantidade de pontos de interesse escolhida para o método foi de 36. Caso o número de pontos utilizado na construção do modelo de forma fosse alterado para valores maiores e que representassem de uma maneira mais completa a forma alvo, os resultados obtidos provavelmente seriam melhores. Entretanto, essa alteração não foi executada para este trabalho devido a quantidade de tempo necessária para um especialista inserir as marcações na base de imagens.

## 7 Conclusões e Trabalhos Futuros

De maneira geral, o método apresentado demonstrou ter melhor desempenho quando o conjunto utilizado para construir o modelo for maior e mais heterogêneo. A quantidade de pontos utilizados para construir o modelo também influi diretamente na capacidade do método de se adequar para as diferentes formas, incrementando o tempo de procura e o consumo de memória pela aplicação.

A vantagem da técnica utilizada neste trabalho é a sua simplicidade e aplicabilidade na utilização conjunta com outros métodos. Para esta situação, os resultados não sobressaíram-se das marcações de pontos de referência manuais, devido aos detalhes mencionados na Seção 6. Porém, realizando as devidas alterações, melhores resultados poderão ser alcançados.

Em trabalhos futuros, o algoritmo proposto será implementado em uma aplicação móvel, voltada para o auxílio de cirurgiões plásticos. O aplicativo realizará a detecção automática de pontos de interesse na face humana e fornecerá ao usuário diversas métricas importantes para o uso em uma análise estética.

## Referências

- Apple (2016). Swift – open source language, <https://swift.org/>. Acessado em: 03-01-2016.
- Bugaighis, I., Mattick, C. R., Tiddeman, B. and Hobson, R. (2014). 3D asymmetry of operated children with oral clefts, *Orthod Craniofac Res* **17**(1): 27–37.
- Cootes, T. F., Edwards, G. J. and Taylor, C. J. (1999). Comparing active shape models with active appearance models, in *Proc. British Machine Vision Conf*, BMVA Press, pp. 173–182.
- Cootes, T. F., Edwards, G. J. and Taylor, C. J. (2001). Active appearance models, *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6): 681–685.
- Cootes, T. F. and Taylor, C. J. (1992). Active shape models, *3rd British Machine Vision Conference*, Springer-Verlag, pp. 266–275.
- Cootes, T. F. and Taylor, C. J. (1997). A mixture model for representing shape variation., in A. F. Clark (ed.), *BMVC*, British Machine Vision Association.
- Cootes, T. F. and Taylor, C. J. (2004). Statistical models of appearance for computer vision, *Technical report*, The University of Manchester School of Medicine.
- Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J. (1995). Active shape models – their training and application, *Computer Vision Image Understood* **61**(1): 38–59.
- Jafri, R. and Arabnia, H. R. (2009). A survey of face recognition techniques., *JIPS* **5**(2): 41–68.
- Kass, M., Witkin, A. and Terzopoulos, D. (1988). Snakes: Active contour models, *International Journal of Computer Vision* **1**(4): 321–331.
- Li, Y. and Ito, W. (2005). Shape parameter optimization for adaboosted active shape model, *10th IEEE International Conference on Computer Vision (ICCV 2005)*, 17-20 October 2005, Beijing, China, pp. 251–258.
- Milborrow, S. and Nicolls, F. (2008). Locating facial features with an extended active shape model, *Proceedings of the 10th European Conference on Computer Vision: Part IV, ECCV '08*, Springer-Verlag, Berlin, Heidelberg, pp. 504–513.
- Rogers, M. and Graham, J. (2002). *Computer Vision – ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part IV*, Springer Berlin Heidelberg, Berlin, Heidelberg, chapter Robust Active Shape Model Search, pp. 517–530.
- Yang, M.-H., Kriegman, D. and Ahuja, N. (2002). Detecting faces in images: a survey, *Pattern Analysis and Machine Intelligence, IEEE Transactions* **24**(1): 34–58.
- Zhou, Y., Gu, L. and Zhang, H.-J. (2003). Bayesian tangent shape model: Estimating shape and pose parameters via bayesian inference, *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'03*, IEEE Computer Society, Washington, DC, USA, pp. 109–116.