

Facial Landmarks Detection on Faulty Datasets with Regression Trees and Principal Component Analysis Parametrization

Abner S. Nascimento, Danilo A. Oliveira,
Maria Raquel L. de Couto, Iális C. de Paula Júnior

Engenharia da Computação – Universidade Federal do Ceará (UFC)
Campus de Sobral – Sobral – CE

abnersousanascimento@alu.ufc.br, daniloreriutaba@gmail.com,
raquelcouto@outlook.com, ialis@sobral.ufc.br

Abstract. Tracking landmarks points of the human face is an essential step for the construction of interfaces capable of taking advantage of the communicative potential of facial expressions. Many strategies based on parametric models and regression algorithms with boosting can be applied to this problem. This paper proposes a solution based on the combined use of principal component analysis and regression trees. The main purpose of the presented method is to reduce the sensibility of the system to the presence of missing labels when trained with faulty datasets, by the adoption of corrective heuristics. On such cases, the proposed model achieves performance similar to the reference results, obtained by training on fault free datasets.

Resumo. A detecção de pontos característicos do rosto humano é essencial para a construção de interfaces capazes de tirar proveito do potencial comunicativo das expressões faciais. Nesse contexto, diversas estratégias baseadas em modelagens paramétricas de formas e algoritmos de regressão com ajuste progressivo podem ser aplicadas. Este trabalho propõe uma solução com uso combinado de modelos baseados em análise de componentes principais e árvores de regressão. A metodologia apresentada tem por objetivo reduzir a sensibilidade à presença de falhas de demarcação nas bases de dados utilizadas para treino, por meio da aplicação de heurísticas corretivas. Em tais casos, o modelo proposto apresenta desempenho similar ao resultado de referência, obtido no treino com amostras íntegras.

1. Introdução

No âmbito da comunicação humana não-verbalizada, a leitura visual de traços faciais representa papel de crítica importância. A expressividade da face torna a externalização de aspectos relacionados ao estado emocional, físico e mental do interlocutor um processo natural e inconsciente, que nem sempre pode ocorrer com a mesma eficiência em outras formas de comunicação [Frith 2009]. Nesse contexto, elementos como a posição das sobrancelhas, lábios e olhos atuam como entidades comunicativas cujo potencial interativo pode se estender também às interfaces digitais. Assim, o uso da face como mecanismo de interação é capaz de prover experiências de uso ricas, que, além de promover intuitividade, dinamicidade e inclusividade, ampliam o espectro de interações oferecidos por modais tradicionais, como teclado e *mouse* [Jaimes e Sebe 2007].

Em essência, a detecção de traços faciais em imagens consiste em determinar a posição de pontos característicos, fixamente demarcados sobre regiões específicas do rosto. Metodologias estatísticas baseadas em modelos paramétricos foram exploradas na literatura para a detecção de pontos característicos de estruturas orgânicas em imagens [Cootes *et al.* 1994]. Nessas abordagens, os perfis de níveis de cinza ao longo de linhas normais ao contorno das formas foram utilizados para calcular o conjunto de parâmetros ótimos. Alternativamente, o cálculo com base na textura da região de interesse também foi estudado em trabalhos posteriores [Cootes *et al.* 2001].

Parametrizações baseadas nos graus de liberdade determinantes na posição dos pontos da forma também foram adotadas para a detecção de poucos elementos característicos da face [Dollár *et al.* 2010]. Para isso, foram empregadas técnicas de regressão, nos quais os pontos alinham-se em suas respectivas localizações após sucessivos ajustes realizados sobre uma estimativa inicial para as variáveis envolvidas.

A estratégia de alinhamento progressivo também foi utilizada em modelos não paramétricos, *i.e.*, algoritmos nos quais o ajuste é feito diretamente sobre a posição dos pontos em vez de parâmetros descritivos [Cao *et al.* 2012]. O ajuste gradual das formas em cada estágio do processo pode ser obtido por meio do trabalho conjunto de árvores de regressão binárias que, embora tenham desempenho pouco satisfatório se aplicadas individualmente, podem alcançar boa acurácia quando usadas em sequência [Kazemi e Sullivan 2014].

Em geral, as técnicas baseadas em ajustes granulares alcançam performances superiores às de modelos paramétricos, uma vez que são capazes de considerar com maior rigor as sutilezas das formas, além de implicarem em custo computacional menor para treino e aplicação. Entretanto, o desempenho de tais abordagens tende a sofrer impacto negativo significativo quando os dados usados para treino apresentam falhas de demarcações, visto que sua capacidade de ajuste depende do correto posicionamento de todos os pontos. Demarcações faltantes podem advir de aspectos comuns à construção de bases para aplicações no mundo real, tais como falha humana, oclusão, ou mesmo imprecisões causadas pela presença de ruído. Nesse cenário, o uso da parametrização, aliado às estratégias de ajuste progressivo das técnicas de regressão, é capaz de produzir modelos menos sensíveis à presença de falhas.

Este trabalho propõe a detecção de pontos faciais característicos por meio da regressão e ajuste progressivo de modelos paramétricos construídos pela análise de componentes principais. Cadeias de árvores treinadas pela diferença entre os níveis de intensidade de pontos da região da face são utilizadas como principal modalidade de regressão. A combinação das duas técnicas é feita com o intuito de incorporar os aspectos positivos de cada uma, bem como reduzir a sensibilidade do algoritmo construído à presença de falhas na base de treino. Ademais, para mitigar os efeitos de tais falhas, propõe-se o uso de uma heurística corretiva baseada na interpolação linear dos pontos ausentes. A metodologia proposta é descrita em detalhes na Seção 2; enquanto os resultados, bem como as discussões sobre eles, são apresentados respectivamente nas Seções 3 e 4.

2. Metodologia

Seja I uma imagem representada como uma matriz bidimensional em que cada elemento contém a intensidade em escala de cinza de um pixel. Para um conjunto de n pontos de in-

teresse de I , dados por $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, e distribuídos ao longo dos traços faciais delimitantes, pode-se definir uma forma S conforme a matriz da Equação (1). Deseja-se, pois, obter a matriz estimativa \hat{S} , de modo que os elementos de \hat{S} se aproximem dos respectivos elementos de S tanto quanto possível.

$$S = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix}. \quad (1)$$

Em síntese, o processo inicia-se pela correção de possíveis falhas na base de dados através da heurística de interpolação linear. Em seguida, o modelo paramétrico é construído com auxílio da análise de componentes principais e uma estimativa inicial é calculada para cada amostra. Por fim, uma cadeia de regressores é treinada para realizar o ajuste dos parâmetros das estimativas iniciais, até que se alcance a convergência.

2.1. Tratamento de pontos faltantes

Quando a base de dados utilizada para a construção do modelo apresenta falhas na demarcação dos pontos, para cada forma do conjunto total existe uma probabilidade ρ , com $0 \leq \rho \leq 1$, de não se conhecer o ponto (x_i, y_i) . Nesses casos, faz-se necessário aplicar heurísticas corretivas de interpolação para preencher as lacunas existentes. Como a análise de componentes principais incorpora variações gerais da forma, em vez de posições específicas dos pontos que a constituem, métodos simples podem ser utilizados, como a interpolação linear.

Seja $K = \langle (x_i, y_i), \dots, (x_j, y_j) \rangle$ uma lista contígua de pontos faltantes entre os índices i e j , ao longo forma S , de modo que (x_{i-1}, y_{i-1}) e (x_{i+1}, y_{i+1}) sejam pontos de extremidade com localizações conhecidas. É possível interpolar os valores dos pontos (x_k, y_k) , com $i \leq k \leq j$, por meio da Equação (2).

$$(x_k, y_k) = (x_{i-1}, y_{i-1}) + \frac{j - k + 1}{j - i + 1} [(x_{j+1}, y_{j+1}) - (x_{i-1}, y_{i-1})]. \quad (2)$$

2.2. Análise de componentes principais

Se todas as formas em um dado conjunto $W = \{S_1, S_2, \dots, S_k\}$ representam consistentemente um mesmo objeto, então seus pontos devem seguir uma distribuição estatística no espaço bidimensional [Cootes *et al.* 1994]. Ao se conhecer tal distribuição, obtém-se um modelo parametrizado capaz de descrever as formas do conjunto, bem como produzir novas variações com aspecto plausível e fiel ao original. A análise de componentes principais – do inglês, *principal component analysis*, ou PCA – é uma abordagem eficiente para determinar essa distribuição. Uma forma S pode ser reconstruída a partir dos parâmetros da PCA conforme a Equação (3).

$$S = \bar{S} + Bp. \quad (3)$$

Na Equação (3), \bar{S} se refere à média das k formas de W , B é uma matriz composta pelos t primeiros autovetores da matriz de covariância das formas e p contém t parâmetros

de deformação. Ao aplicar-se variações nos elementos do vetor \mathbf{p} , é possível produzir formas que não existiam no conjunto original.

Antes do cálculo do modelo é preciso normalizar as formas de W . Para normalizar uma forma \mathbf{S} , com n pontos (x_i, y_i) , efetua-se a translação, escalonamento e rotação de \mathbf{S} pelos respectivos fatores τ , λ e θ descritos na Equação (4) [Kendall 1989]. Os termos (z_i, w_i) são pontos de \mathbf{S}_1 , após sua normalização por τ e λ .

$$\tau = \sum_{i=1}^n \frac{(x_i, y_i)}{n}; \quad \lambda = \sqrt{\sum_{i=1}^n \frac{x_i^2 + y_i^2}{n}}; \quad \theta = \arctan \left(\frac{\sum_{i=1}^n y_i z_i - x_i w_i}{\sum_{i=1}^n x_i z_i + y_i w_i} \right). \quad (4)$$

2.3. Aplicação das árvores de regressão

O cálculo da estimativa $\hat{\mathbf{S}}$ de uma forma \mathbf{S} é feito com base em uma cadeia de regressores constituídos por múltiplas árvores de regressão. Nos regressores, cada árvore atua como um estimador fraco, capaz de utilizar o resíduo da árvore anterior para fazer uma predição rudimentar do ajuste que precisa ser realizado. Assim, a aplicação sucessiva de estimadores fracos reduz gradativamente o resíduo, até que se alcance a convergência próxima do resultado esperado [Hill e Lewicki 2006].

A atuação dos regressores consiste, pois, essencialmente, em aplicar sucessivas correções ao posicionamento dos pontos de uma estimativa inicial. Portanto, o processo deve iniciar-se pelo cálculo de tal estimativa. Posteriormente, cada árvore do regressor deve dividir o conjunto de treino em agrupamentos, de modo que cada grupo contenha formas que sofrerão ajustes similares. Por conseguinte, a predição de novas amostras deve ser realizada pela aplicação dos mesmos critérios de divisão empregados na construção da árvore.

2.3.1. Estimativa inicial

Para aplicar as árvores de regressão, é preciso definir, a princípio, uma estimativa inicial para \mathbf{S} . A forma média $\bar{\mathbf{S}}$ serve bem a essa finalidade, porque permite alcançar, já de início, maior proximidade com o resultado final do que uma estimativa aleatória, por exemplo. Além disso, essa abordagem facilita a aplicação da Equação (3), pois $\bar{\mathbf{S}}$ é o primeiro termo da soma. No segundo termo, a matriz \mathbf{B} provém do cálculo da análise de componentes principais e o vetor de parâmetros \mathbf{p} deverá ser fornecido pela cadeia de regressores. A forma média é, então, posicionada na região central da face, detectada com o auxílio do método de histograma de gradientes orientados [Dalal e Triggs 2005].

Após o cálculo da estimativa inicial, tanto $\hat{\mathbf{S}}$ quanto \mathbf{S} são normalizadas pelos fatores τ , λ e θ da Equação (4). Assim, é possível aplicar sobre elas a análise de componentes principais, para calcular os parâmetros $\hat{\mathbf{p}}$ e \mathbf{p} . Logo, a mudança necessária sobre $\hat{\mathbf{p}}$ para obter \mathbf{p} , *i.e.*, o resíduo da estimativa inicial, é dada por $\Delta\mathbf{p} = \mathbf{p} - \hat{\mathbf{p}}$. Deseja-se, pois, que as árvores de regressão sejam capazes de fornecer um ajuste aproximado $\Delta\hat{\mathbf{p}}$ para diminuir o resíduo tanto quanto possível.

2.3.2. Processamento das amostras

Estruturalmente, as árvores de regressão são árvores binárias nas quais os nós internos dividem o conjunto de amostras com base em critérios específicos aplicados sobre os dados de treino. Nós de um mesmo nível aplicam o mesmo critério. Para incorporar a informação dos pixels da imagem I no processo de treino, um conjunto $M = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_q\}$ de pontos de referência aleatórios da região da face é associado a cada regressor. A partir deles, os critérios de divisão das árvores são definidos como tuplas $\mu = (\mathbf{m}_i, \mathbf{m}_j, \alpha)$, com $1 \leq i, j \leq q$ e $i \neq j$. O valor α é um fator limiar de diferença entre as intensidades dos pixels em \mathbf{m}_i e \mathbf{m}_j . Tanto α quanto o primeiro ponto, \mathbf{m}_i , são sorteados aleatoriamente com distribuição de probabilidade uniforme. Entretanto, para a escolha de \mathbf{m}_j , são favorecidos pontos com distância euclidiana mais próxima de \mathbf{m}_i .

Portanto, em cada nó β da árvore, para cada imagem dos dados de treino, se a intensidade do pixel de coordenada \mathbf{m}_i subtraída da intensidade do pixel de coordenada \mathbf{m}_j superar o limiar α , a imagem é encaminhada para a subárvore direita β_{dir} . Caso contrário, é encaminhada à subárvore esquerda β_{esq} . Ao fim das sucessivas divisões, a predição final $\Delta\hat{\mathbf{p}}$ da árvore é determinada pela média dos vetores $\Delta\mathbf{p}$ das amostras remanescentes em cada folha.

A predição de uma árvore pode ser melhorada significativamente se os critérios de divisão para cada nível forem escolhidos dentre os melhores de um conjunto $C = \{\mu_1, \mu_2, \dots, \mu_c\}$, gerado de forma randômica. Um critério pode ser avaliado quanto à sua qualidade pela função $f : C \rightarrow \mathbb{R}$, definida conforme a Equação (5), na qual $\beta_{\mu, \text{dir}}$ e $\beta_{\mu, \text{esq}}$ referem-se ao conjunto de imagens que foram encaminhadas para a subárvore direita e esquerda pelo critério μ .

$$f(\mu) = \|\beta_{\mu, \text{dir}}\| \sum_{\forall \Delta\mathbf{p} \in \beta_{\mu, \text{dir}}} \Delta\mathbf{p} + \|\beta_{\mu, \text{esq}}\| \sum_{\forall \Delta\mathbf{p} \in \beta_{\mu, \text{esq}}} \Delta\mathbf{p}. \quad (5)$$

2.3.3. Treino e predição

No processo de treino, cada árvore recebe uma lista de tuplas dada por $D = \langle (\mathbf{I}_1, \mathbf{p}_1, \hat{\mathbf{p}}_1, \Delta\mathbf{p}, M_1), \dots, (\mathbf{I}_k, \mathbf{p}_k, \hat{\mathbf{p}}_k, \Delta\mathbf{p}, M_k) \rangle$, cujos termos correspondem a cada uma das k amostras de imagens com formas \mathcal{S} demarcadas. Uma vez que o treinamento de uma árvore é concluído, o processo para efetuar a predição de uma nova amostra consiste em aplicar os critérios μ à imagem em cada nível até que se atinja um nó folha. Em seguida, a predição $\Delta\hat{\mathbf{p}}$ que foi associada ao nó durante a aprendizagem é multiplicada a um fator de amortecimento ϕ , e somada aos parâmetros \mathbf{p} da amostra correspondente, de modo a obter um novo vetor de parâmetros atualizado \mathbf{p}' . Pode-se, então atualizar a forma da estimativa $\hat{\mathcal{S}}$ através da Equação (3).

Após a aplicação de sucessivas árvores de regressão, a estimativa $\hat{\mathcal{S}}$ deve convergir para uma configuração que, embora mais próxima de \mathcal{S} do que a princípio, ainda requer ajustes. Faz-se necessário, portanto, atualizar o conjunto de pontos de referência M das futuras árvores, para que reflitam a mudança de aspecto de $\hat{\mathcal{S}}$ e uma nova cadeia de ajustes possa ocorrer. Seja $\hat{\mathcal{S}}'$ a forma $\hat{\mathcal{S}}$ antes do ajuste pelas árvores. Para que sejam reposicionados, os pontos \mathbf{m} são rotacionados por um ângulo γ em torno do ponto mais

próximo de \hat{S}' com o raio da rotação escalonado conforme um fator de escalonamento ψ . Os fatores γ e ψ são, respectivamente, a rotação e escala que transformam \hat{S}' em \hat{S} [Kazemi e Sullivan 2014].

2.4. Avaliação da estimativa

Para averiguar o grau de aproximação entre a predição e a forma real, o erro é calculado como a média das distâncias euclidianas entre os pontos de S e seus correspondentes em \hat{S} , como descrito na Equação (6). Para mitigar diferenças de resolução das imagens das amostras, as distâncias são normalizadas pela distância interocular d , calculada a partir da diferença entre os pontos médios dos trechos correspondentes aos olhos em S [Zhu e Ramanan 2012].

$$E(S, \hat{S}) = \frac{1}{n} \sum_{i=1}^n \frac{\sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}}{d}. \quad (6)$$

3. Resultados

A metodologia descrita foi aplicada sobre a base de dados Helen Dataset [Le *et al.* 2012]. Na base, há 2330 imagens demarcadas com 194 pontos distribuídos ao longo dos olhos, nariz, lábios, sobrancelhas e maxilar. Do conjunto total de dados, 2000 amostras foram usadas para treino dos modelos, 315 para testes e 15 foram descartadas por apresentarem pontos demarcados fora do espaço definido pela região de coordenadas positivas. As imagens foram processadas com espaço de cores em escala de cinza.

A princípio, o método foi testado com todas as formas definidas integralmente, isto é, sem pontos faltantes. A fim de comparar a influência do número de parâmetros da análise de componentes principais no desempenho das árvores de regressão, diferentes quantidades de parâmetros foram testadas. O erro resultante para as configurações utilizadas está exposto na Tabela 1. Nas Figuras 1(a) a 1(c) é possível comparar o ajuste realizado pelos modelos treinados com 80, 100 e 120 parâmetros na PCA em relação à forma real, *i.e.*, o ajuste ótimo, indicado na Figura 1(d).

Tabela 1. Comparativo entre os erros resultantes nas amostras de teste após a aplicação dos regressores treinados pelo método proposto.

80 parâmetros	100 parâmetros	120 parâmetros
0,1922	0,1098	0,1090

De modo a avaliar o desempenho da técnica quando aplicada em bases com falhas na demarcação dos pontos, foram construídas bases de treino deficitárias com 40%, 60% e 80% de pontos faltantes. Para isso, pontos selecionados aleatoriamente ao longo das curvas que compõem as formas de treino originais foram removidos, com exceção das extremidades. Assim, para correção das lacunas, a heurística de interpolação linear da Equação (2) pode ser utilizada. O resultado produzido por esse procedimento sobre um exemplar das amostras encontra-se ilustrado na Figura 2. É possível observar que a interpolação ocasiona perda na suavidade das curvas, além de distorções que aumentam com a quantidade de pontos removidos.

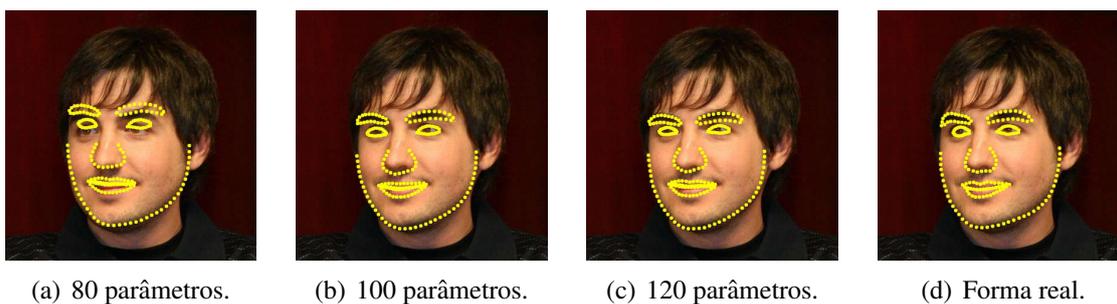


Figura 1. Resultado de predições feitas pelo método proposto sobre uma amostra do Helen Dataset, com diferentes quantidades de parâmetros da PCA, em comparação à demarcação real dos pontos.

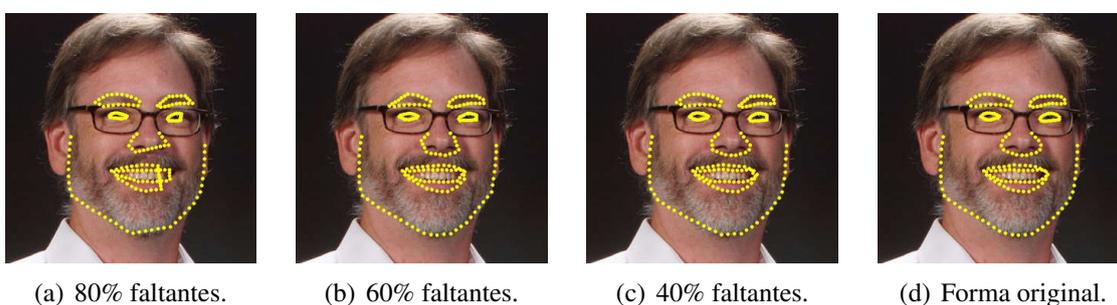
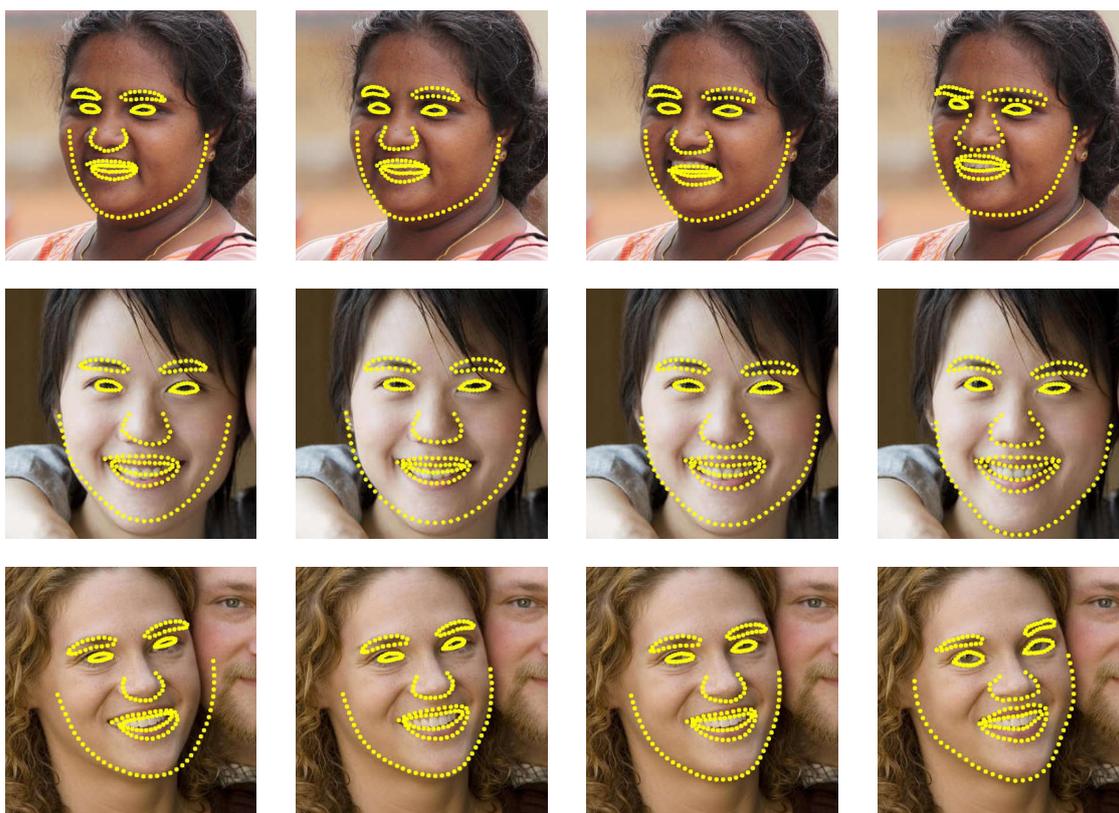


Figura 2. Amostras de treino com pontos faltantes, corrigidos pela heurística de interpolação linear, em comparação com a forma original.

Sobre as bases faltantes, foram treinadas cadeias de regressores a partir de modelos PCA com diferentes quantidades de parâmetros. Os erros restantes após a aplicação dos modelos nas amostras de teste estão dispostos na Tabela 2. As diferenças percentuais entre os modelos treinados com amostras que sofreram remoção de pontos e os modelos treinados com as amostras completas encontram-se na Tabela 3. As taxas obtidas permitem constatar a proximidade entre os resultados alcançados pelos modelos faltantes e os modelos de referência, treinados sobre a base de dados completa. Na Figura 3, são apresentados exemplos visuais dos resultados obtidos sobre as amostras de teste. Apesar das deformações resultantes da interpolação linear, os ajustes puderam conservar a suavidade das curvas e consistência das formas, mesmo na ausência de 80% dos pontos.

No intuito de verificar a redução progressiva do erro ao longo da cadeia de regressores, os modelos de 120 parâmetros treinados com as amostras faltantes foram avaliados em cada estágio do processo de teste. Os resultados estão expostos nos gráficos da Figura 4. Observa-se que a proximidade entre os modelos treinados com bases deficitárias e o modelo treinado com a base completa é consistente ao longo de todo o processo de treinamento.



(a) 80% faltantes. (b) 60% faltantes. (c) 40% faltantes. (d) Demarcação real.

Figura 3. Exemplos de resultados obtidos a partir de modelos treinados com pontos faltantes em comparação com as respectivas demarcações reais. Foram utilizados 120 parâmetros da PCA.

Tabela 2. Erro resultante nas amostras de teste após a aplicação dos regressores treinados em amostras com pontos faltantes corrigidas por interpolação linear.

Pontos faltantes	Quantidade de parâmetros da PCA		
	80	100	120
40%	0,1926	0,1132	0,1133
60%	0,1929	0,1152	0,1140
80%	0,1930	0,1138	0,1166

Tabela 3. Diferença percentual entre os erros dos modelos treinados em amostras com pontos faltantes e os modelos treinados com amostras completas.

Pontos faltantes	Quantidade de parâmetros da PCA		
	80	100	120
40%	+0,20%	+3,08%	+3,96%
60%	+0,38%	+4,93%	+4,60%
80%	+0,41%	+3,63%	+7,02%

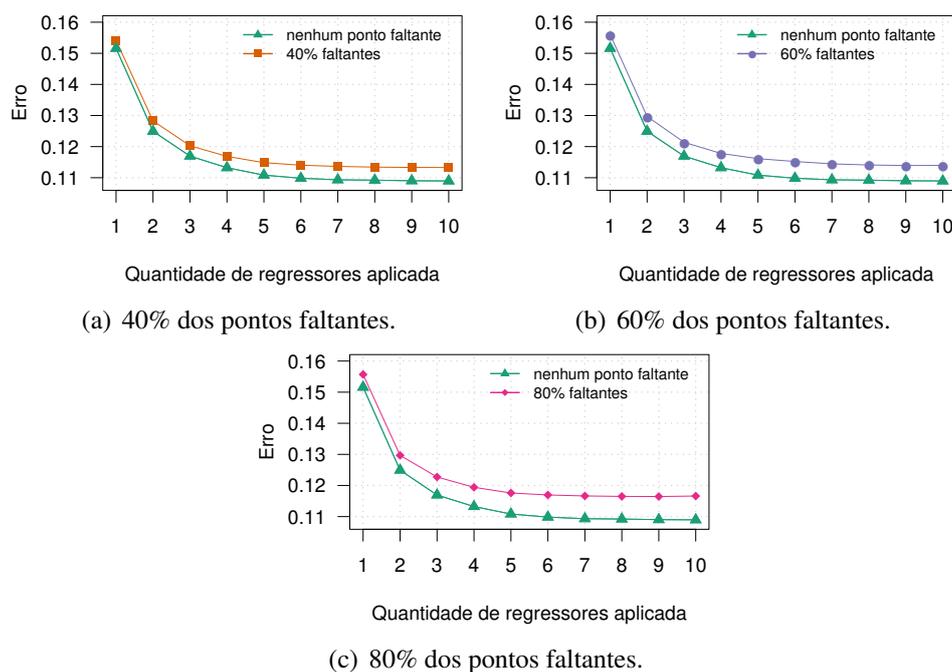


Figura 4. Redução do erro em função do número de regressores para modelos treinados com parte dos pontos faltantes em comparação à mesma métrica para o modelo treinado integralmente. Foram usados 120 parâmetros da PCA.

4. Discussões

Por meio dos resultados obtidos após a aplicação da metodologia com variações na quantidade de parâmetros PCA, constata-se que o erro tende a diminuir quando a quantidade de parâmetros aumenta. De fato, mais parâmetros implicam em maior grau de variabilidade na representação dos dados pela análise de componentes principais. Consequentemente, o modelo torna-se capaz de se ajustar melhor às sutilezas das formas. Todavia, em bases com muitos pontos faltantes, o erro do modelo com 120 parâmetros mostrou-se ligeiramente maior. Essa diferença se deve ao superajuste do modelo com excesso de parâmetros às discontinuidades que foram imprecisamente corrigidas pela interpolação linear.

Ao longo de toda a cadeia de regressores, o desempenho alcançado pela metodologia proposta sobre os dados faltantes mostrou-se próximo do que foi obtido com as amostras íntegras. Desse modo, embora simples, a heurística corretiva de interpolação linear é suficiente para que o modelo generalizado pela PCA possa ajustar-se com níveis de precisão similares aos obtidos com as formas completas. Tal flexibilidade existe porque a análise de componentes principais permite explorar a distribuição estatística da posição dos pontos no espaço para realizar o ajuste. Assim, o efeito das correções imprecisas pode ser reduzido, desde que a quantidade de parâmetros utilizadas na PCA não seja excessiva.

Ademais, o uso de estratégias de interpolação para corrigir possíveis faltas permite maior proveito das amostras de treino, ainda que parte das demarcações não estejam bem determinadas. Mais imagens à disposição para o treino trazem maior diversidade para as amostras. Assim, as árvores de regressão tornam-se capazes de fornecer previsões adequadas em um escopo maior de iluminações, cenários e outras variações circunstanciais das imagens sobre as quais serão aplicadas.

5. Conclusão

A combinação da estratégia paramétrica baseada na análise de componentes principais e o ajuste sequencial da cadeia de árvores de regressão resulta em um modelo de detecção de pontos faciais menos sensível à ausência de demarcações nas amostras de treino. A simples interpolação linear dos pontos faltantes mostrou-se suficiente para que os regressores pudessem realizar o ajuste com desempenho próximo do que obtiveram quando treinados sobre as bases de dados com formas íntegras. Desse modo, pelo uso da heurística de correção, mesmo as amostras falhas podem ser aproveitadas no processo de treino. Assim, conforme proposto, a metodologia baseada na combinação das abordagens supracitadas permite localizar pontos faciais característicos com desempenho satisfatório, mesmo quando os dados disponíveis apresentam falhas.

Referências

- Cao, X., Wei, Y., Wen, F., e Sun, J. (2012). Face alignment by explicit shape regression. Em *2012 IEEE Conference on Computer Vision and Pattern Recognition*, páginas 2887–2894.
- Cootes, T. F., Edwards, G. J., e Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):681–685.
- Cootes, T. F., Hill, A., Taylor, C. J., e Haslam, J. (1994). Use of active shape models for locating structures in medical images. *Image and vision computing*, 12(6):355–366.
- Dalal, N. e Triggs, B. (2005). Histograms of oriented gradients for human detection. Em *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, páginas 886–893. IEEE.
- Dollár, P., Welinder, P., e Perona, P. (2010). Cascaded pose regression. Em *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, páginas 1078–1085. IEEE.
- Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical transactions of the royal society of London B: Biological sciences*, 364(1535):3453–3458.
- Hill, T. e Lewicki, P. (2006). *Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry, and Data Mining*.
- Jaimes, A. e Sebe, N. (2007). Multimodal human–computer interaction: A survey. *Computer vision and image understanding*, 108(1-2):116–134.
- Kazemi, V. e Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. Em *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, páginas 1867–1874.
- Kendall, D. G. (1989). A survey of the statistical theory of shape. *Statistical Science*, 4(2):87–99.
- Le, V., Brandt, J., Lin, Z., Bourdev, L., e Huang, T. S. (2012). Interactive facial feature localization. Em *European Conference on Computer Vision*, páginas 679–692. Springer.
- Zhu, X. e Ramanan, D. (2012). Face detection, pose estimation, and landmark localization in the wild. Em *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, páginas 2879–2886. IEEE.