

Banco de dados de faces 3D

– Relatório Técnico –

Jesús P. Mena-Chalco¹, Roberto M. Cesar Junior¹, Luiz Velho²

¹Instituto de Matemática e Estatística

Universidade de São Paulo

São Paulo - Brasil

²Instituto de Matemática Pura e Aplicada

Rio de Janeiro - Brasil

Janeiro de 2011

Resumo

Tradicionalmente as expressões faciais humanas têm sido exploradas usando fotografias 2D ou sequências de vídeo 2D. Análises e sínteses baseadas em faces 2D são de difícil realização para dados com grandes variações de pose e variações sutis de expressão facial. Os bancos de dados de faces 3D existentes geralmente incluem faces com variação de pose e com poucas aquisições de expressões faciais que permitam a investigação de mudanças de expressões espontâneas inerentemente da natureza humana. Neste relatório, são descritos brevemente os principais bancos de dados de faces apresentados na literatura até o presente momento, dando especial atenção a conjuntos de dados 3D com expressões faciais.

1 Introdução

O entendimento sobre expressões faciais é de grande relevância, desde análises psicológicas (Russell e Dols, 1997) até a criação de modelos evolutivos da expressão facial humana (Schmidt e Cohn, 2001). Uma vez que a expressão facial é reconhecida, esta informação pode ser usada para ajudar a pessoa identificada e/ou adaptar automaticamente a interface de interação, por exemplo, através de uma interface homem-máquina (Pantic e Rothkrantz, 2003). Note que, bancos de dados contendo amostras de expressões faciais 3D são fundamentais para estudo e avaliação de métodos de reconhecimento de faces sob diferentes condições, registro de modelos faciais, correção de pose, análise e síntese de faces e de expressões faciais. Em particular, a face neutra é usada geralmente para o reconhecimento de faces. Entretanto, a face neutra de uma pessoa pode: (i) não casar à face da mesma pessoa com expressão facial diferente, ou (ii) pode casar com a face neutra de uma outra pessoa (Faltemier *et al.*, 2007).

(*) O conteúdo do presente relatório técnico é de única responsabilidade dos autores.

Diversos são os bancos de dados 3D idealizados para o reconhecimento de faces e ocasionalmente usados para a análise de expressões faciais. A maioria dos bancos de dados disponíveis, no entanto, contém uma quantidade limitada de expressões faciais, variações de pose, e registro de dados da forma geométrica com a informação de textura, ou de oclusão. Nesta seção são listados os bancos de dados de faces 3D mais representativos. Salientamos que o conjunto de dados IMPA-FACE3D, criado para manter informação de textura e geometria registrada para todas as expressões faciais, foi desenvolvido no mesmo período da publicação do conjunto Bosphorus (Savran *et al.*, 2008). Este último conjunto de dados é, até agora, o mais representativo na literatura por sua completude.

Aqui devemos reforçar que conjuntos de dados comuns e metodologias de avaliação comuns, ainda não foram definidas (Yin *et al.*, 2006), e são necessárias para comparar, validar, resolver e estender problemas relacionados a análises e sínteses de expressões faciais 3D (similar aos propostos no desafio FRGC (Phillips *et al.*, 2005) e na metodologia de avaliação FERET Phillips *et al.* (2000)).

2 Banco de dados disponíveis

A seguir descrevemos as características mais relevantes dos conjuntos de dados publicamente disponíveis. Para uma avaliação rápida, veja na Tabela 1 um resumo com as informações mais relevantes dos conjuntos de dados de faces 3D.

Bosphoru (Savran *et al.*, 2008): É um conjunto de dados que inclui aquisições de 105 indivíduos com diferentes tipos de expressões, variação de pose e diferentes tipos de oclusão. Ao todo, foram considerados 60 homens e 45 mulheres, sendo que a maioria dos indivíduos estão entre os 25 e 35 anos de idade. Foram adquiridos entre 31 e 54 amostras para todos os indivíduos, sumando 4652 amostras no total. Atualmente, este é o conjunto de dados conhecido, mais completo e representativo no estado-da-arte.

Neste conjunto, foram consideradas aquisições correspondentes a dois tipos de expressões faciais: (i) expressões faciais baseadas em 28 unidades de ação¹; e (ii) expressões faciais referidas às seis expressões universais entre as raças humanas, proposto por Ekman e Friesen (1971). Em complemento às expressões faciais, foram consideradas 14 aquisições com diferentes poses e oclusões nas regiões dos olhos, boca e testa. De forma automática foram eliminados os ruídos nos dados, entretanto, não foi realizado nenhum processamento para tratar os problemas de auto-occlusões nem de descontinuidades (buracos), comumente presentes nas regiões de boca e lados do nariz.

IMPA-FACE3D: (Mena-Chalco *et al.*, 2008) Este conjunto de dados inclui aquisições de 38 indivíduos com uma amostra de face neutra, seis amostras correspondentes a expressões faciais universais e outras 5 amostras referentes a expressões contendo boca e olhos abertos

¹As unidades de ação (*actions units*) do sistema de codificação de ação facial (FACS, *Facial action coding system*), proposta por Ekman e Friesen (1978), foram idealizadas para construir ou categorizar sistematicamente blocos de expressões faciais correspondente a emoções humanas. Cada unidade de ação está relacionada com a ativação de diversos conjuntos de músculos da face. Assim, por meio destas pode-se rotular quase qualquer expressão facial anatomicamente possível.

e/ou fechados. Também foram consideradas 2 amostras correspondentes aos perfis laterais dos indivíduos. Ao todo, o conjunto de dados está composto por 22 homens e 16 mulheres, sendo que a maioria dos indivíduos estão na faixa etária entre 20 e 50 anos. Foram adquiridas 14 amostras para todos os indivíduos, sumarizando 532 amostras no total.

Este banco de dados foi criado no IMPA para a análise de expressões faciais e a reconstrução de faces 3D. Não são consideradas aquisições com oclusão de objetos (e.g. óculos), nem com variações de pose. Foram considerados procedimentos para eliminar os ruídos originados usualmente pelas reflexões especulares dos aparelhos de iluminação na face. Adicionalmente, os problemas de auto-occlusão e de descontinuidades (buracos) são tratados com técnicas de processamento de imagens.

York-3DFace: (**Heseltine et al., 2008**) Este conjunto de dados foi criado por pesquisadores da Universidade de York e é composto por 5250 aquisições correspondentes a 350 indivíduos. Foram consideradas 15 aquisições para cada pessoa, contendo 5 tipos de expressões diferentes, 6 orientações de pose, e 4 distâncias de aquisição diferentes.

Os dados foram adquiridos por uma unidade de visão estéreo, desenvolvida pelos autores, provida de 2 câmeras calibradas para obtenção da forma de geometria, e uma câmera para captura da textura facial. Para 1770 faces deste conjunto de dados, foram realizadas operações de normalização ou alinhamento de pose por meio da detecção automática de pontos característicos faciais.

ND-2006: (**Faltemier et al., 2007**) É um conjunto de dados criado por pesquisadores da universidade de Notre Dame e está composta de 13450 amostras correspondentes a 888 indivíduos. Para cada indivíduo foram adquiridas até 63 amostras usando o escâner 3D Vivid 910 da Minolta. No conjunto foi considerada a face neutra e as expressões faciais correspondentes a alegria, tristeza, surpresa, desgosto e alternativamente outras expressões faciais. Salientamos que nem todas essas expressões foram adquiridas para todas as pessoas. As aquisições da face neutra pertencem a mais de um quarto do conjunto de dados.

CASIA 3D: (**Zhong et al., 2007**) Este conjunto de dados, ainda em crescimento, está sendo mantido pelo Centro de Pesquisas em Biometria e Segurança do Instituto de Automação da *Chinese Academy of Sciences* (CASIA). Atualmente, o banco de dados é composto por 4624 aquisições correspondentes a 37 ou 38 amostras de 123 indivíduos. Foram consideradas, adicionalmente, aquisições de faces, expressões com diversas variações de pose, e cinco tipos diferentes de iluminação. O equipamento usado é o escâner 3D Vivid 910 da Minolta.

BU-3DFE: (**Yin et al., 2006**) Este é um conjunto de dados criado pela Universidade de Binghamton e especialmente projetado para a análise da face neutra e das seis expressões faciais universais, proposta por **Ekman e Friesen (1971)**, cada uma com 4 níveis de intensidade na expressão: baixa, média, alta, e exagerada. Ao todo foram 2500 aquisições correspondentes a 25 amostras de 100 indivíduos (44 homens e 56 mulheres com idades entre 18 e 70 anos). Não foram consideradas variações de pose nem de iluminação.

Os dados foram pré-processados para eliminar as regiões externas da faces como por exemplo pescoço e ombros. Adicionalmente aos dados de informação geométrica e de textura,

foi realizada uma correção da pose, e também associados 83 pontos característicos faciais (posicionados manualmente) em cada forma facial 3D.

FRAV3D: (Conde, 2006) É um conjunto de dados criado pela Universidade Rey Juan Carlos, e foi criada para sua aplicação em reconhecimento de faces 3D. Ao todo foram consideradas 1696 aquisições correspondentes a 16 amostras para cada indivíduo. Foram adquiridas 4 faces neutras, uma de sorriso, uma com boca aberta, 8 com a face neutra e poses diferentes. Também foram consideradas duas aquisições com iluminação não controlada. É importante destacar que este banco de dados não apresenta o registro entre a informação de textura e a forma da face, i.e., a informação de textura não esta relacionada à informação de geometria.

BJUT-3D-R1: (Beijing University of Technology, 2005) Este é um banco de dados criado pela Universidade de Tecnologia de Beijing e é usado para análise de faces 3D com especial foco na aquisição de indivíduos de raça oriental. Foi usado um ambiente com iluminação controlada e um escâner 3D CyberWare 3030RGB/PS como equipamento para adquirir a face completa (360 graus) em todas as amostras (Cyberware, 2010).

Ao todo, o banco está conformado por 2000 aquisições correspondentes a 4 amostras de 500 indivíduos (250 homens e 250 mulheres com idades entre 16 e 49 anos). As amostras correspondem à face neutra e as expressões de alegria, raiva e surpresa. Não foram consideradas variações de pose nem de iluminação.

Para a aquisição, todos os indivíduos vestiram uma touca (azul ou vermelha) para prender o cabelo e deixar exposta a região da face, pescoço e orelhas. Os dados foram pré-processados para eliminar os possíveis ruídos e regiões correspondentes ao cabelo e pescoço.

FRGC 2.0: (Phillips *et al.*, 2005) É um conjunto de dados idealizado pelo Instituto Nacional de Padrões e Tecnologia, em parceria com a Universidade de Notre Dame, e é comumente usado para avaliar sistemas de reconhecimento de faces 2D e 3D. Para o desafio *Face Recognition Grand Challenge*, foram criados seis experimentos para medir o desempenho dos reconhecedores computacionais de faces considerando imagens com dois tipos de iluminação (controlada e sem controle), associados a partições de treinamento e de validação.

Atualmente, o banco é composto por 4007 aquisições 3D correspondentes de 1-22 amostras de 466 indivíduos com idades não informadas. Foi usado o escâner 3D Vivid 910 da Minolta para adquirir todas as formas faciais com suas correspondentes informações de textura. Não foram consideradas variações de pose nas aquisições 3D.

GavabDB: (Moreno e Sánchez, 2004) Este é um conjunto de dados criado pela Universidade *Rey Juan Carlos*, é composto por 549 aquisições correspondentes a 61 indivíduos (45 homens e 16 mulheres caucasianos com idades compreendidas entre 18 e 40 anos). Para cada indivíduo foram obtidas 9 diferentes aquisições da forma geométrica sem textura.

No conjunto de dados aparecem variações sistemáticas de pose e expressão facial dos indivíduos. Em particular, existem 2 aquisições frontais e neutras, 4 aquisições com leves rotações e sem expressões faciais, e 3 aquisições frontais que apresentam diferentes expressões faciais como sorriso, riso e alguma outra expressão aleatória escolhida pelo indivíduo (não existe, neste conjunto, uma padronização das expressões faciais).

3D_RMA: (Beumier e Acheroy, 2001) É um conjunto de dados criado pela *Royal Militar Academy* da Bélgica e é composta por 720 aquisições correspondentes a 120 indivíduos (106 homens e 14 mulheres). Para cada indivíduo foram obtidas 6 amostras de faces neutras e com variações de poses diferentes . Foram também consideradas aquisições de indivíduos com barba e bigode. O sistema de aquisição 3D foi construído com uma câmera e um projetor e está baseado em uma codificação por luz estruturada². A informação de textura não foi associada as amostras adquiridas.

²Os sistemas de aquisição por luz estruturada utilizam um par projetor-câmera para iluminar a cena com um padrão de luz estruturada. O sistema ótico do projetor tem um papel similar ao de uma segunda câmera nas medições estereoscópicas e o padrão projetado facilita a correspondência entre os pontos (Carvalho *et al.*, 2005).

Tabela 1: Bancos de dados de faces 3D. ‘Ind.’ refere-se ao número de indivíduos. ‘Ams.’ refere-se ao número de amostras por indivíduos. ‘Total’ refere-se ao número total de amostras no conjunto de dados.

Nome	Ind.	Ams.	Total	Expressões faciais	URL
Bosphorus (Savran <i>et al.</i> , 2008)	105	31-54	4652	face neutra, 6 expressões universais e outras expressões correspondentes a 28 unidades de ação.	bosphorus.ee.boun.edu.tr Bogazici University, Turquia.
IMPA-FACE3D (Mena-Chalco <i>et al.</i> , 2008)	38	14	532	face neutra, 6 expressões universais, e 5 expressões exageradas	www.vision.ime.usp.br/~jmena/IMPA-FACE3D IME/USP & IMPA, Brasil.
York-3DFace (Heseltine <i>et al.</i> , 2008)	350	15	5 250	face neutra, alegria, raiva, olhos fechados, e sobranceiras levantadas.	www.cs.york.ac.uk/arch University of York, Reino Unido.
ND-2006 (Faltennier <i>et al.</i> , 2007)	888	1-63	13 450	face neutra, alegria, tristeza, surpresa, desgosto, e outras diversas.	www.nd.edu/~cvr1 University of Notre Dame, Estados Unidos.
CASIA-3D (Zhong <i>et al.</i> , 2007)	123	37-38	4 624	sorriso, riso, raiva, surpresa e olhos fechados.	www.cbsr.ia.ac.cn/english/3DFace%20Databases.asp Institute of Automation chinese Academy of Sciences, China.
BU-3DFE (Yin <i>et al.</i> , 2006)	100	25	2 500	face neutra, 6 expressões universais (cada uma com 4 intensidades)	www.cs.binghamton.edu/~lijun/Research/3DFE Binghamton University, Estados Unidos.
FRAV3D (Conde, 2006)	106	16	1 696	face neutra, sorriso, e boca aberta.	www.frav.es/research/facerecognition/FRAV3D Rey Juan Carlos University, Espanha.
BJUT-3D-R1 (Beijing University of Techno 2005)	500	4	2 000	face neutra, alegria, raiva e surpresa.	www.bjut.edu.cn/sci/multimedia/mul-1lab/3dface/face_database.htm Beijing University of Technology, China.
FRGC2.0 (Phillips <i>et al.</i> , 2005)	466	1-22	4 007	face neutra, alegria, tristeza, surpresa, e cansaço (<i>puffy</i>).	www.frvt.org/FRGC National Institute of Standards and Technology, Estados Unidos.
GavabDB (Moreno e Sánchez, 2004)	61	9	549	face neutra, sorriso, riso, e alguma outra aleatória.	gavab.esct.urjc.es/recursos.html Universidad Rey Juan Carlos, Espanha.
3D_RMA (Beumier e Acheroy, 2001)	120	6	720	face neutra.	www.sic.rma.ac.be/~beumier/DB/3d_rma.html Royal Military Academy, Bélgica.

3 Considerações finais

Nos últimos anos, aplicações sobre faces humanas 3D passaram a ter grande popularidade devido à sua relativa vantagem sobre suas correspondentes representações 2D. Esta vantagem refere-se à independência de iluminação e independência de pontos de vista da câmera (*viewpoint*) sob dados de geometria (Savran *et al.*, 2008). Essa tendência também deve-se à facilidade de obtenção de dados geométricos e uso de escâners 3D.

Finalmente, é importante destacar que existe uma lista relacionada a bancos de dados de faces 2D e 3D, continuamente atualizada e disponível publicamente em www.face-rec.org/databases. Esta página, mantida por Grgic e Delac (2010), serve como ponto inicial para pesquisadores na prospecção de conhecimento sobre faces humanas e em particular o reconhecimento computacional de faces.

Referências

- Beijing University of Technology(2005)** Beijing University of Technology. The BJUT-3D large-scale chinese face database. Technical report, Beijing University of Technology. Technical Report of The Multimedia and Intelligent Software Technology Beijing Municipal Key Laborator.
- Beumier e Acheroy(2001)** C. Beumier e M. Acheroy. Face verification from 3D and grey level clues. *Pattern Recognition Letters*, 22(12):1321–1329.
- Carvalho *et al.*(2005)** P.C. Carvalho, L. Velho, A. Sá, E. Medeiros, A.A. Montengro, A. Peixoto, e L.A.R. Escriba. *Fotografia 3D. 25 Colóquio Brasileiro de Matemática*. IMPA.
- Conde(2006)** C. Conde. *Biometría: Reconocimiento facial mediante fusión 2D y 3D*. Dykinson SL, Madrid.
- Cyberware(2010)** Cyberware. Cyberware 3D scanner homepage. <http://www.cyberware.com/products/scanners/>, 2010. Último acesso em 21/09/2010.
- Ekman e Friesen(1978)** P. Ekman e W. V. Friesen. The facial action coding system: A technique for the measurement of facial movement. Em *Consulting Psychologists*, Palo Alto, California.
- Ekman e Friesen(1971)** P. Ekman e W.V. Friesen. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2):124–129.
- Faltemier *et al.*(2007)** T. Faltemier, K. Bowyer, e P. Flynn. Using a multi-instance enrollment representation to improve 3D face recognition. páginas 1–6.
- Grgic e Delac(2010)** M. Grgic e K. Delac. Face recognition homepage. <http://www.face-rec.org>, 2010. Último acesso em 17/09/2010.

- Heseltine et al.(2008)** T. Heseltine, N. Pears, e J. Austin. Three-dimensional face recognition using combinations of surface feature map subspace components. volume 26, páginas 382–396. Elsevier.
- Mena-Chalco et al.(2008)** J.P. Mena-Chalco, R.M. Cesar-Jr, e L. Velho. Banco de dados de faces 3D: IMPA-FACE3D. Technical report, Instituto de Matemática Pura e Aplicada - IMPA - VISGRAF Laboratory, Rio de Janeiro, RJ, Brazil. TR01.
- Moreno e Sánchez(2004)** A.B. Moreno e A. Sánchez. GavabDB: a 3D face database. Em *Workshop on Biometrics on the Internet*, páginas 77–85, Vigo.
- Pantic e Rothkrantz(2003)** M. Pantic e L.J.M. Rothkrantz. Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, 91(9):1370–1390.
- Phillips et al.(2000)** P.J. Phillips, H. Moon, P. Rauss, e S.A. Rizvi. The feret evaluation methodology for face-recognition algorithms. *IEEE transactions on pattern analysis and machine intelligence*, 22(10):1090–1104.
- Phillips et al.(2005)** P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, e W. Worek. Overview of the face recognition grand challenge. Em *IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2005*, páginas 947–954.
- Russell e Dols(1997)** J.A. Russell e J.M.F. Dols. *The Psychology of Facial Expression*. Cambridge University Press.
- Savran et al.(2008)** A. Savran, N. Alyüz, H. Dibeklioglu, O. Celiktutan, B. Gökberk, B. Sankur, e L. Akarun. Bosphorus database for 3D face analysis. *Biometrics and Identity Management*, páginas 47–56.
- Schmidt e Cohn(2001)** K.L. Schmidt e J.F. Cohn. Human facial expressions as adaptations: Evolutionary questions in facial expression research. *American Journal of Physical Anthropology*, 116(S33):3–24.
- Yin et al.(2006)** L.J. Yin, X.Z. Wei, Y. Sun, J. Wang, e M.J. Rosato. A 3D facial expression database for facial behavior research. Em *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, páginas 211–216.
- Zhong et al.(2007)** C. Zhong, Z. Sun, e T. Tan. Robust 3D face recognition using learned visual codebook. Em *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR*, páginas 1–6.

