



Experimentação para a Construção de uma Base de Imagens com Registro de Emoção

Pedro Vieira de Carvalho*, Sara S. de Oliveira, Murilo M. Pechoto, Paula D. P. Costa, Léo Pini Magalhães

Resumo

Neste trabalho, foi construída uma base de imagens de expressões faciais espontâneas rotuladas por emoção. A base é composta por vídeos RGB-D (com informação de cor e profundidade) e seus respectivos rótulos, e será utilizada no treinamento de algoritmos de classificação para identificação automática de emoções em expressões faciais.

Palavras-chave:

Reconhecimento de emoções, processamento de faces, processamento de imagem

Introdução

Os estudos deste trabalho fazem parte da área de Computação Afetiva, e contribuirão para o desenvolvimento de sistemas capazes de interpretar o estado emocional de usuários, provendo as respostas mais adequadas às suas emoções.

O objetivo geral do projeto de iniciação científica é desenvolver um sistema capaz de identificar expressões faciais de usuários que interagem com um computador a partir de uma câmera RGB-D posicionada frontalmente.

Neste contexto, os objetivos específicos do presente trabalho são a construção de uma base de imagens tridimensionais de expressões faciais rotuladas e explorar algoritmos de classificação utilizando essa base.

Resultados e Discussão

O trabalho é dividido nas seguintes etapas: desenvolvimento do software de captura, gravação dos vídeos de expressões faciais, extração de pontos característicos (*landmarks*), treinamento de algoritmos de classificação e análise dos resultados.

Para obter os vídeos com expressões faciais espontâneas, o sistema exibe na tela do computador uma série de vídeos curtos para o voluntário. Simultaneamente, ele aciona a câmera Intel RealSense F200 e registra em vídeo 3D as expressões faciais do participante. A Figura 1 mostra a montagem do ambiente de captura.

Além disso, depois de cada vídeo, o participante responde a algumas perguntas referentes à emoção que sentiu e dessa maneira consegue-se catalogar as expressões faciais gravadas de acordo com as respostas do participante, e assim rotular os elementos da base.

Com a conclusão do software de aquisição, deu-se início à etapa de captura dos vídeos. Para isso, contou-se com a colaboração de 43 voluntários, com os quais foram gravados mais de 1200 vídeos de reações espontâneas rotuladas.

A partir dos vídeos extrai-se 78 pontos característicos (*landmarks*) da face em cada quadro, como indicado na Figura 1. Esses pontos são identificados utilizando o kit de desenvolvimento da Intel RealSense¹.

Com isso, para cada vídeo gravado há 78 pontos tridimensionais por quadro, rotulados segundo a emoção que aquele vídeo representa. Esses serão os dados utilizados para a última etapa do trabalho. Os próximos passos previstos são a extração de características (*features*) geométricas² a partir de *landmarks* e posterior treinamento de algoritmos de classificação.

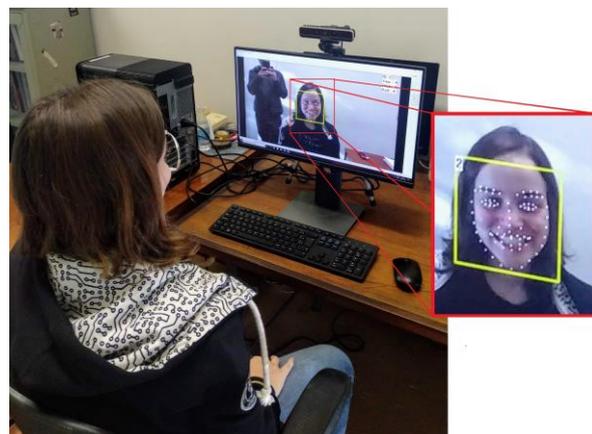


Figura 1. Montagem do ambiente de captura com destaque para *landmarks* identificados na face

Conclusões

O trabalho permite o estudo de um fluxo típico de aprendizado de máquina, desde a construção da base, extração de características, treinamento de modelo, classificação e análise.

Por fim, a criação da base de vídeos 3D abre um leque de possibilidades de trabalhos futuros, como por exemplo avaliar o impacto da coordenada de profundidade para a identificação de emoções em imagens de expressões faciais.

Agradecimentos

Este trabalho foi apoiado pelo PIBIC (Programa Institucional de Bolsas de Iniciação Científica), programa do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) em parceria com a Unicamp.

¹ Intel® RealSense™ SDK 2016 R2 Documentation. Disponível em: <https://software.intel.com/sites/landingpage/realsense/camera-sdk/v1.1/documentation/html/index.html?doc_devguide_whats_new_in_sdk.htm>. Acesso em: 7 fev. 2019.

² LI, Xiaoli; RUAN, Qiuqi; MING, Yue. 3D facial expression recognition based on basic geometric features. In: **IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS**. IEEE, 2010. p. 1366-1369.