



## RECONHECIMENTO FACIAL EM IMAGENS CAPTURADAS POR CÂMERAS DE SEGURANÇA UTILIZANDO EIGENFACES E UMA VERSÃO MODIFICADA DO LBP

Fernando Beletti\*<sup>1</sup>, Josué Silva de Morais<sup>1</sup>

<sup>1</sup>FEELT – Universidade Federal de Uberlândia

**Resumo** - Este documento descreve a utilização da base de dados SCface para reconhecimento facial em imagens capturadas por câmeras de segurança. Esta é uma base de dados interessante, pois provê dados que simulam bem situações de segurança reais, em que câmeras de segurança estão espalhadas por todos os lugares. Buscando solucionar o problema de identificação das faces em baixa resolução optou-se por utilizar a técnica de eigenfaces e também a utilização da transformada census modificada, não como ferramenta de detecção de faces, mas sim como ferramenta de pré-processamento para extração de características. Além disso, discute-se a dificuldade relacionada ao processo de reconhecimento quando há uma galeria pequena e principalmente à diferença gigantesca de qualidade entre as imagens de treino e teste.

**Palavras-Chave** - análise principal de componentes, eigenfaces, LBP, reconhecimento fácil, SCface

### FACE RECOGNITION WITH IMAGES ACQUIRED FROM SURVEILLANCE CAMERAS USING EIGENFACES AND A MODIFIED VERSION OF LBP

**Abstract** - This paper describes the use of the SCface database for face recognition on images captured by surveillance cameras. This database is interesting because it provides data that mimics real security situations, where surveillance cameras are spread everywhere. In order to solve the low-resolution face identification problem, the eigenfaces technique was chosen and, also a modified version of the Census Transform, which was not used as a face detector, but as a feature extractor from the faces. Besides, it is discussed the issues with performing facial recognition in conditions where there is a small gallery and mostly when there is a huge contrast between training and test datasets resolution.

**Keywords** - eigenfaces, facial recognition, LBP, principal component analysis, SCface

\*fbeletti@ufu.br

### I. INTRODUÇÃO

O reconhecimento facial é uma tarefa amplamente realizada por um ser humano no seu dia a dia. De maneira intrínseca e até mesmo surpreendente, um humano é capaz de reconhecer um rosto familiar, mesmo diante de inúmeras adversidades, por exemplo, obstáculos, diferentes ângulos e condições adversas de iluminação. Buscando replicar a capacidade de detecção e reconhecimento, não só de faces, pelos humanos, surgiram inúmeros trabalhos no campo da visão computacional que buscam automatizar esses processos através do aprendizado de máquina e técnicas como reconhecimento de padrões.

Estudos relacionados ao reconhecimento facial automatizado existem desde o século passado, ainda na década de 60, quando já se pensava em sua utilização como ferramenta de cooperação com a força policial para o reconhecimento de indivíduos infringem as leis [1, 2].

Com o passar dos anos o *hardware* foi se aprimorando e o ato de processar uma grande quantidade de dados deixou de ser um problema e tornou-se um facilitador para os estudos nesta área, o que é possível perceber com o crescimento da quantidade de publicações na área concomitante com o crescimento da capacidade de processamento dos computadores. Não só artigos, mas inúmeros sistemas comerciais foram desenvolvidos e distribuídos, sendo a autenticação de usuários nos celulares um exemplo de um sistema amplamente e diariamente utilizado de forma corriqueira.

De modo a testar a robustez e também auxiliar no desenvolvimento destes sistemas de reconhecimento facial é necessário a utilização de conjuntos de dados do mundo real. Sendo assim, existem vários conjuntos de dados públicos, em sua grande maioria adquiridos em estudos realizados em universidades ao redor do mundo, em que são possíveis encontrar faces em várias condições para que sejam utilizadas pelos sistemas. Neste trabalho utilizou-se a base de dados SCface, proveniente da Universidade de Zagreb [3].

Esta base de dados é bastante particular, por se tratar de imagens capturadas por câmeras de segurança, criando um problema pouco explorado atualmente, em que as imagens de prova são de baixa resolução e as imagens de treinamento são de alta resolução.

Desta forma, este trabalho tem por propósito realizar o reconhecimento facial utilizando a base de dados citada, com um processamento diferente do utilizado pelos donos da base de dados, sendo neste utilizada uma versão modificada dos padrões locais binários (mLBP).

Nas próximas seções a composição da base de dados será melhor apresentada, além dos métodos utilizados e também os resultados obtidos.

## II. BASE DE DADOS E ALGORITMOS

### A. A base de dados SCface

Poucos estudos utilizam esta base de dados por se tratar de uma base de dados desafiadora [4]. Esta base de dados contém um conjunto de imagens de 130 indivíduos, capturadas por 5 câmeras de segurança durante o dia, 2 câmeras em modo noturno e 1 em modo infravermelho, sendo que todas as câmeras tiraram fotos dos indivíduos em 3 distâncias diferentes. Além destas imagens cada indivíduo teve registrada por uma câmera de alta definição uma foto, em ambiente controlado, de seus rostos.

Exemplos de imagens adquiridas podem ser vistas na Figura 1. Maiores detalhes de como as imagens foram adquiridas são apresentados na referência [3].

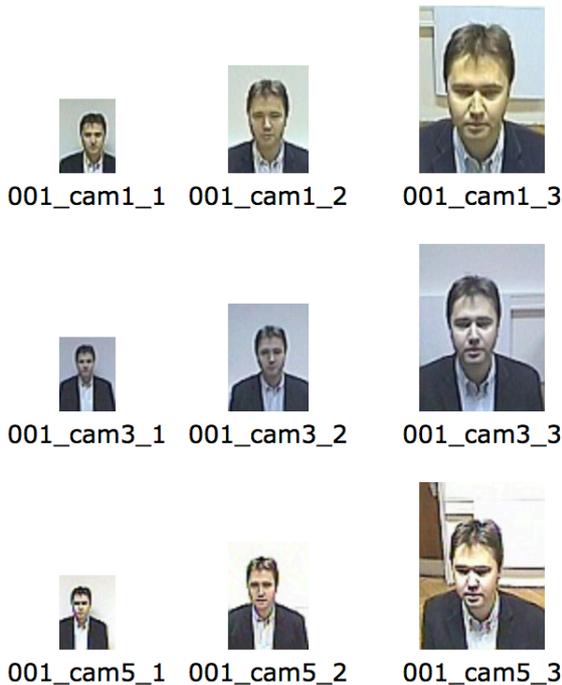


Figura 1: Exemplo imagens adquiridas [4]

### B. A versão modificada dos padrões locais binários (mLBP)

A transformada LBP, também conhecida como Transformada Census, originalmente surgiu para descrever a textura das imagens tratadas. Nos últimos anos foi utilizada para compensar e normalizar a iluminação nos algoritmos de reconhecimento facial. [6]

Esta transformada é baseada no mapeamento dos pixels de uma janela através da comparação da intensidade do pixel central com a dos pixels da sua vizinhança, sendo geralmente esta janela de 3x3.

Seja  $N(x)$  uma vizinhança espacial local de um dado pixel central  $x$  tal que  $x \notin N(x)$ . A Transformada Census gera uma cadeia de bits representando quais pixels em  $N(x)$  possuem intensidade de brilho menor que  $I(x)$ . Considerando que as intensidades de brilho são sempre zero ou positivas, uma definição formal pode ser dada da seguinte maneira:

Assumindo que a função de comparação  $h(I(x), I(y))$  retorna 1 se  $I(x) < I(y)$ , e 0 senão, e tomando  $\cup$  como sendo o operador de concatenação, a Transformação Census em  $x$  é definida como:

$$LBP(x) = \cup_{y \in N(x)} h(I(x), I(y)) \quad (1)$$

Um esquema da transformada pode ser visto na Figura 2.

Tomando que a janela utilizada seja 3x3 a cadeia de bits resultante poderá apresentar  $2^8 = 256$  possíveis valores [5]. Já na transformada modificada, o pixel central é considerado, esta modificação é proposta para corrigir alguns casos em que o LBP não é capaz de capturar corretamente a padrão local de uma imagem [6].

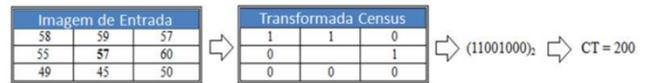


Figura 2: Exemplo janela de comparação

Mas obviamente não podemos comparar a intensidade do pixel central com ela mesma, então na transformada modificada há uma nova vizinhança  $N'(x) = N(x) \cup I(x_c)$  tal que agora as comparações são realizadas em relação à média das intensidades da vizinhança  $h'(I(x), I(y))$ .

$$mLBP(x) = \cup_{y \in N'(x)} h'(I(x), I(y)) \quad (2)$$

Neste caso as comparações resultam em  $2^9 - 1 = 511$  possíveis valores, representando melhor o padrão local da imagem.

### C. Algoritmo de reconhecimento facial

No presente trabalho o método de reconhecimento facial utilizado são as *eigenfaces*, uma das técnicas primordiais para realizar tal tarefa [1].

Este método de reconhecimento facial consiste em projetar linearmente todas as imagens que estão sendo comparadas em um espaço dimensional reduzido através de uma transformada obtida por Análise Principal de Componentes (PCA), que não é foco deste trabalho. Entretanto, de maneira simples, as imagens iniciais são representadas em um novo espaço com dimensão menor representados por autovalores e autovetores.

O principal objetivo da redução de dimensão é tornar o problema mais simples, pois assim o calcula das distâncias entre as figuras comparadas são menores. Por exemplo uma figura 32x32 pixels representa um espaço com 1024 dimensões, sendo realizadas, portanto, 1024 comparações de vértices para se determinar a proximidade entre as faces e enfim definir a identificação.

Nos experimentos realizados seguiu-se a linha básica de reconhecimento e identificação de faces e segue os

protocolos previstos em [3]. Primeiramente as faces são encontradas nas imagens, posteriormente pré-processadas, por fim o conjunto passa pelo processo de PCA para então serem identificadas [7].

Apesar de este algoritmo ser considerado ultrapassado para a resolução de alguns problemas, optou-se por utilizá-la para efeitos de comparação com os resultados obtidos pelos pesquisadores responsáveis pela base de dados, uma vez que os mesmos também utilizaram eigenfaces.

### III. RESULTADOS DOS EXPERIMENTOS

Nos experimentos realizados primeiramente as imagens foram transformadas em escala de cinza. Depois foram detectados os principais pontos da face e alinhadas os olhos com uma operação de translação e rotação das imagens. Posteriormente as imagens foram reduzidas (ou aumentadas) para uma única resolução, no caso 64x64 pixels. Neste caso a dimensão do problema é 4096, entretanto para simplificar as comparações e tornar o algoritmo mais rápido, utilizou-se o PCA com as 130 primeiras componentes, conseguindo uma representação de 100% das características das imagens originais, isto é, o problema foi diminuído para 130 dimensões, que representa aproximadamente 3% da complexidade inicial do problema de identificação das faces.

É possível acompanhar as alterações realizadas nas imagens do mesmo indivíduo na Figura 3 (alta resolução) e Figura 4 (baixa resolução).

Em um primeiro momento as imagens são convertidas para a escala de cinza de modo a reduzir a dimensão do problema e acelerar o processamento, que antes se tratava de 3 canais (padrão RGB) e agora é tratada apenas em um canal.

As faces são alinhadas e cortadas todas conforme um único padrão para que durante o processo de reconhecimento, estruturas similares estejam em posições similares, tornando o processo mais simples e efetivo. Por exemplo, os olhos e boca tendem a estar sempre na mesma posição, assim quando calculada a distância entre as faces, comparam-se regiões próximas. O corte é realizado para que fique em foco apenas a face, eliminando ao máximo o fundo, que não deve ser parte do algoritmo de treinamento, nem teste, do reconhecimento facial.

Motivado por [6], optou-se por utilizar mLBP sobre as imagens com o intuito simplesmente de atuar como um pré-processador, não como um método de reconhecimento facial. Utilizou-se a distância euclidiana para se encontrar as faces mais próximas e então identifica-las. Os resultados foram extremamente ruins, ao contrário do que se esperava de acordo com a literatura, em que esta combinação (mLBP e distância euclidiana) apresentaram os melhores números. Sendo corretos apenas 4,5% dos resultados.

Diante de tal resultado, ainda motivado pelo mesmo trabalho, utilizou-se a distância de Hamming adaptada para o problema de 2 pontos, sendo os resultados diferentes para as imagens de algumas câmeras, mas no geral inexpressíveis resultados em 4,8% de respostas corretas.

Apesar dos resultados, é importante ressaltar que em [6], o pré-processamento mLBP é utilizado para reduzir problemas de iluminação nas imagens adquiridas, mas no conjunto de dados SCface este não é um problema, o que comprova que este pré-processamento não adicionou melhoria os resultados

obtidos inicialmente em [3] que representam 4,7% de acertos.



Figura 3: Mudanças nas imagens originais de acordo com o *pipeline* de processamento. 1- aquisição das imagens, 2- escala de cinza, 3- alinhamento e corte, 4- pré-processamento 5- identificação



Figura 4: Mudanças nas imagens de baixa resolução de acordo com o *pipeline* de processamento.

Avaliando os resultados obtidos, optou-se por utilizar outros métodos clássicos de pré-processamento de imagens, que são a normalização das intensidades dos pixels, e a equalização dos histogramas das imagens. Métodos amplamente utilizados e básicos para auxiliar nas tarefas de reconhecimento facial [4, 5, 7], basicamente servem para realçar estruturas e características nas imagens de modo que quando aplicado a análise principal dos componentes, sejam obtidos componentes que representem as imagens de maneira melhor.

Neste novo experimento, notou-se melhoria nos resultados, sendo que neste momento acertou-se 8,2% das identificações.

Tabela 1: Resultados obtidos

Método	Acertos
mLBP	4,8%
<b>Método próprio</b>	<b>8,2%</b>
[4]	4,7%

### IV. CONCLUSÕES

O conjunto de dados SCface é composto por uma galeria com fotos únicas de cada indivíduo adquiridas em altíssima qualidade, que devem ser comparadas com imagens de baixíssima qualidade dos mesmos indivíduos. Desta forma surgem dois grandes problemas ainda pouco explorados no campo do reconhecimento facial, o primeiro, o problema da

galeria pequena, onde existem poucos, no caso um único exemplo de cada indivíduo, o que dificulta o reconhecimento facial; e o segundo, o problema de se comparar imagens de alta resolução com imagens de baixa resolução.

Geralmente, os sistemas de reconhecimento facial comuns, pautados em autenticação, funcionam em ambientes controlados, em que o indivíduo a ser autorizado contém mais de uma imagem de exemplo, e, principalmente a câmera que captura a imagem de cadastro, é a mesma que captura a imagem de teste. O fato de se utilizar câmeras diferentes para aquisição das imagens de teste e imagens de treino é um grande vilão na base de dados utilizada.

Em um primeiro momento os resultados obtidos aparentam ser extremamente pobres, devido às baixas taxas de acerto. Contudo, estes resultados reiteram como esta base de dados é extremamente desafiadora e importante para desenvolvimento de sistemas mais robustos.

Os resultados não são tão ruins se comparados aos resultados obtidos na referência, sendo que em casos especiais (câmeras mais bem posicionadas em que as faces foram capturadas mais de frente à câmera) na literatura original os resultados não foram superiores a 8% e nos testes realizados foram alcançados 12%.

Conclui-se que utilizar da Transformada Census Modificada não trás grandes melhorias ao processo de reconhecimento facial para esse conjunto de dados se comparado aos métodos tradicionais de pré-processamento de imagens, provavelmente devido à grande diferença de qualidade entre as imagens de treino e teste. Para se obter melhores resultados em trabalhos futuros, devem ser utilizados outros métodos de pré-processamento como métodos de super resolução capazes de tratar as imagens de péssima resolução em resolução melhor ou em um espaço dimensional intermediário.

Enfim, a base de dados utilizada é desafiadora e bastante desoladora se utilizados métodos convencionais para a resolução de seus problemas. Contudo é um dos melhores conjuntos de dados faciais disponíveis na internet que proporciona faces nessas condições, que são perfeitas para simular o trabalho da polícia moderna, em que os presos são fichados e retratados por imagens de alta definição geralmente únicas, por exemplo fotos de passaporte ou CNH, e uma vez soltos ou em fuga podem ser monitorados por câmeras comuns de segurança que auxiliadas por um sistema robusto de detecção nessas condições podem se tornar a realidade, com baixo custo e precisão razoável se comparado

aos próprios humanos, que em condições de baixa resolução também apresentam dificuldades de reconhecimento e identificação de pessoas, conforme é possível observar na Figura 4.

## AGRADECIMENTOS

Este projeto foi financiado pela CAPES. Partes deste trabalho usam a base de dados SCface. Créditos são dados à Faculdade de Engenharia Elétrica e Computação da Universidade de Zagreb por disponibilizarem esse conjunto de imagens.

## REFERÊNCIAS

- [1] W. K. Taylor, "Machine Learning and recognition of faces", in *Electronics Letters*, vol. 3, no. 9, pp. 436-437, September 1967. doi: 10.1049/el:19670340
- [2] A. J. Goldstein, L. D. Harmon and A. B. Lesk, "Identification of human faces," in *Proceedings of the IEEE*, vol. 59, no. 5, pp. 748-760, May 1971. doi: 10.1109/PROC.1971.8254
- [3] Mislav Grgic, Kresimir Delac, Sonja Grgic, SCface - surveillance cameras face database, *Multimedia Tools and Applications Journal*, Vol. 51, No. 3, February 2011, pp. 863-879
- [4] Nguyen, Huu-Tuan & Caplier, Alice. (2012). Elliptical Local Binary Patterns for Face Recognition. 7728. 10.1007/978-3-642-37410-4\_8.
- [5] Ribeiro, Ícaro & Marana, Aparecido & Chiacchia, Giovani. (2011). Reconhecimento de Faces Sob Diferentes Condições de Iluminação Utilizando PCA e a Transformada Census. *Learning and Nonlinear Models*. 9. 138-144. 10.21528/LNLM-vol9-no2-art6.
- [6] Javier Ruiz-del-Solar, Julio Quinteros, Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches, *Pattern Recognition Letters*, Volume 29, Issue 14, 2008, Pages 1966-1979, ISSN 0167-8655, <https://doi.org/10.1016/j.patrec.2008.06.015>.
- [7] H. Nguyen and A. Caplier, "Local Patterns of Gradients for Face Recognition," in *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 8, pp. 1739-1751, Aug. 2015. doi: 10.1109/TIFS.2015.2426144