

UM ROBUSTO RECONHECIMENTO FACIAL POR FILTRO DE GABOR CURVO E ENTROPIA

Eucassio Gonçalves Lima Júnior, Luis Henrique Silva,
Cornelia Janayna Pereira Passarinho, Ricardo de Andrade Lira Rabêlo
Universidade Federal do Piauí (UFPI) - Teresina, Piauí - Brasil
eucassiojr@gmail.com, lhvogado@gmail.com
janaynapassarinho@gmail.com, ricardoalr@ufpi.edu.br

Resumo: Reconhecimento facial é um desafio na área de processamento de imagens, considerando-se que as imagens sofrem influência de diversos fatores tais como: posição da cabeça, iluminação, expressão facial e oclusão. Este trabalho apresenta uma abordagem para reconhecimento facial, caracterizando-se pela robustez a oclusão, variação de iluminação e pela baixa dimensionalidade do vetor de características. A abordagem baseia-se na combinação do filtro de Gabor curvo e entropia de forma a manter os dados mais representativos da imagem em um vetor de características de dimensionalidade menor do que aquele utilizando o Gabor original. A abordagem proposta foi avaliada em diversos cenários: variações de iluminação, oclusão por óculos de sol, variação de iluminação com oclusão por óculos de sol, oclusão por cachecol, variação de iluminação com oclusão por cachecol, mudança de expressão facial e todos os cenários juntos. Os resultados obtidos por meio dos experimentos superam as abordagens estado da arte disponíveis na literatura em 3 dos 5 testes (iluminação, iluminação com oclusão por óculos de sol e iluminação com oclusão por cachecol) e adiciona outros 2 testes: mudança de expressão facial e todos os cenários juntos.

Palavras-chaves: Entropia; Filtro de gabor; Oclusão; Reconhecimento facial; Redução de dimensionalidade.

Abstract: Face recognition is a challenge in image processing area, considering that the images are influenced by various factors such as head position, lighting, facial expression and occlusion. This paper presents an approach to facial recognition, characterized by robustness occlusion, lighting variation and low dimensionality of the feature vector. The approach is based on a combination of curved and entropy Gabor filter in order to maintain the most representative image data in vector dimensionality characteristics lower than that using the original Gabor. The proposed approach was evaluated in several scenarios: lighting changes, occlusion by sunglasses, change lighting occlusion by sunglasses, occlusion scarf, lighting variation occlusion scarf, twisted face and all the scenarios together. The results obtained through experiments outweigh approaches state of the art available in the literature in 3 of the 5 tests (lighting, occlusion by sunglasses and lighting occlusion scarf) and adds other 2 tests: change of facial expression and all scenarios together.

Keywords: Gabor filter; Entropy; Dimensionality reduction; Occlusion; Facial recognition.

I. INTRODUÇÃO

O reconhecimento facial tornou-se nos últimos anos um campo de pesquisa que desenvolveu-se rapidamente [1]. Apresentando-se como um dos campos mais populares quanto ao processamento e análise de imagens, a principal motivação é a necessidade do desenvolvimento de diversos sistemas que necessitam dessa tecnologia. Dentre eles podemos citar:

sistemas que permitam a identificação de pessoas para acesso a dados restritos, sistemas de controle de acesso e segurança, monitoramento de multidões, localização de fichas criminais, interatividade controlada por expressões faciais, dentre outros.

Os algoritmos de reconhecimento facial podem ser classificados em 3 grandes grupos: tradicionais, heterogêneos e homogêneos [2]. Os algoritmos tradicionais (Figura 1a) utilizam uma única imagem inicial e um algoritmo, os heterogêneos (Figura 1b) aplicam n algoritmos diferentes usando n diferentes imagens de entrada e finalmente os homogêneos (Figura 1c) usam o mesmo algoritmo com n parâmetros diferentes e uma mesma imagem de entrada. A abordagem proposta nesse trabalho, pode ser classificada como homogênea, extraindo o máximo de um algoritmo, variando os seus parâmetros.

O reconhecimento facial tem por objetivo primordial identificar uma pessoa baseada em uma imagem de sua face. Para isso é necessário comparar as imagens iniciais de busca, com n amostras de um banco de dados facial. Atualmente existem diversas fontes de imagens de face, já armazenadas para os mais variados fins, tais como: grandes bancos de dados de fotografias de passaportes, carteiras de motoristas, identidades, dentre outros relatados em [3]. Apesar dos avanços ocorridos nas últimas décadas fruto de intensas pesquisas, alguns desafios permanecem como problemas não resolvidos [4]. Resultantes principalmente de condições não controladas do ambiente, tais como oclusão, posicionamento, variações de expressões faciais e mudanças de iluminação, esses desafios continuam sendo motivo de pesquisa em reconhecimento facial.

A abordagem proposta tem por base o uso do filtro de gabor curvo [2] para a extração do vetor de características de uma imagem, para em seguida ser aplicada a entropia. Com o objetivo de manter os dados mais representativos da imagem em um vetor de dimensionalidade menor, utilizou-se a entropia. Por fim, essa representação computacional da imagem é fornecida como entrada ao SVM (Support Vector Machines) [5], responsável por efetuar a classificação das imagens. Para a validação da abordagem proposta, comparou-se os resultados obtidos com outros de abordagens que possuem o mesmo objetivo de reconhecer faces com em ambiente com oclusão, são elas: SRC[6], LRC[7], CESR[8], FSS[9], SGLasso[10] e DCC[11].

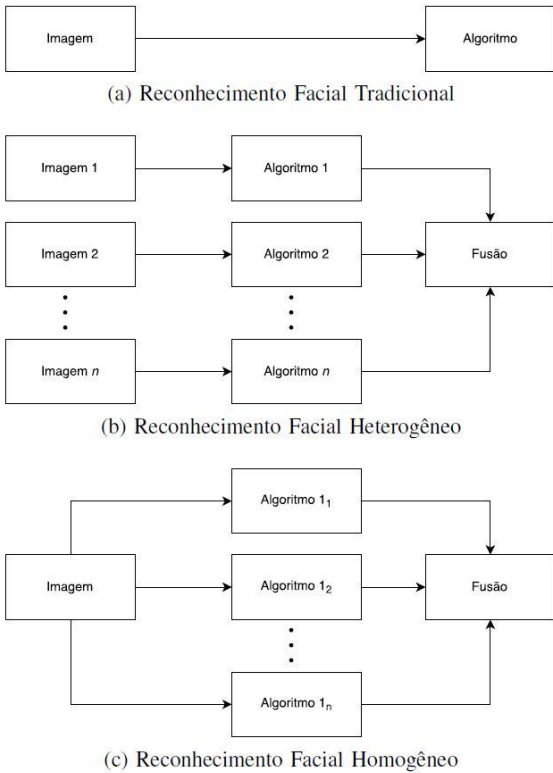


Figura 1- Categorias de métodos de reconhecimento facial

Um aspecto de um sistema robusto de reconhecimento facial é a representação das características de uma imagem através de um vetor com baixa dimensionalidade. Atributos como armazenamento, tempo de recuperação e identificação devem ser otimizados nesse tipo de sistema. A técnica de representação geralmente inicia-se com um procedimento de redução de dimensionalidade. Devido ao fato de que a alta dimensionalidade de uma imagem original dificulta ou mesmo inviabiliza o classificador de obter bons resultados [12].

Conforme proposto inicialmente em [13], a medida que aumenta-se o número de atributos que devem ser considerados, cresce exponencialmente o número de atributos do conjunto de dados. Para um algoritmo de aprendizagem de máquina, definir um modelo preciso de predição com diversos atributos torna-se mais complexo. Esse fenômeno é chamado de maldição da dimensionalidade, por ele o número de exemplos necessários para se aprender um certo conceito cresce exponencialmente de acordo com o número de atributos [14]. Usualmente, o desempenho de um classificador tende a diminuir a partir de uma determinada quantidade de atributos [15], apesar desses serem significativos ao problema.

A etapa final da abordagem de reconhecimento facial é a classificação. Seu objetivo é identificar os indivíduos por meio de um algoritmo de aprendizagem de máquina. A execução do algoritmo é precedida da realização do treinamento, definindo um modelo preciso de predição que dada uma nova imagem é capaz de identificar o indivíduo.

As principais contribuições deste trabalho são relacionadas abaixo:

1) A inclusão e adequação do banco de filtro de gabor curvo para reconhecimento facial em conjunto com a entropia como técnica de redução de dimensionalidade. Alcançou-se significativos avanços em recentes experimentos de reconhecimento facial.

2) Desenho de uma abordagem robusta baseada no uso do filtro de gabor curvo, entropia [16] e SVM para reconhecimento facial.

As demais seções deste artigo estão organizadas da seguinte forma: na Seção II, é detalhado o referencial técnico exibindo uma visão geral do funcionamento do banco de filtro de Gabor e sua variação curva, Entropia e SVM. A Seção III formula o problema de reconhecimento facial que está sendo abordado e estabelece a solução proposta nesse trabalho. Nessa seção são apresentados os detalhes técnicos de implementação e aplicação do banco de filtros de Gabor, Entropia e SVM. Os resultados dos experimentos efetuados são mostrados e discutidos na Seção IV. Por último, na Seção V são apresentadas as conclusões do artigo.

A. Trabalhos relacionados

As abordagens SRC (Sparse Representation-based Classification) e CESR (Correntropy-based sparse representation) são baseadas em representação esparsa. Nessas abordagens, ao invés de amostrar todas as características da imagem, torna-se suficiente adquirir apenas alguns pedaços de informação e de maneira aleatória. Apesar de não procurar por pontos específicos esse tipo de representação de imagem é suficiente para reconstruir a imagem.

Em [6] é apresentado o problema de reconhecimento facial em imagens, destacando os principais desafios. Entre eles, as distorções ocasionadas por variações de iluminação, expressões posicionamento e oclusão que ocorrem e trazem incerteza na representação da imagem. Como abordagem proposta para solucionar tais problemas é desenvolvido um novo modelo de regressão (DCC) que provê medidas de erro multiescala e compactness e colaboração interclasses. A abordagem promove um significativo avanço nos resultados dos experimentos com diversas bases de imagens. Entretanto, verifica-se que os valores obtidos de acurácia para faces com oclusão mostram-se abaixo de outros trabalhos, como o apresentado neste artigo.

Feature-sign search em [9], apresenta o problema do uso de sparse coding em grandes bases de imagens devido o alto custo computacional. No artigo é desenvolvido um novo algoritmo que de forma iterativa resolve 2 problemas de otimização convexa, na descoberta dos melhores vetores bases para a representação das imagens. A abordagem apresenta-se robusta para problemas de oclusão em blocos contínuos localizados na parte inferior da face. Observa-se o valor de acurácia de 95,38% no reconhecimento de indivíduos na base de imagens Ar Face para oclusão por cachecol. Entretanto, a apresenta dificuldades sobretudo ao tratar a oclusão sobre a região dos olhos.

Para lidar com problemas de oclusão e de corrupção no reconhecimento facial é apresentado em [8] um algoritmo chamado CESR (Correntropy-based sparse representation). Nele é combinado representação esparsa não negativa e o critério da máxima correntropy.

Os demais trabalhos que funcionam como comparativo da abordagem proposta, fundamentam-se em regressão linear.

LRC (Linear Regression Classification), FSS (Feature-sign search), SGLasso (Structured graphical lasso) e DCC (Discriminative and compact coding) consideram a imagem de teste como uma combinação linear de algumas das imagens de treinamento.

Dentre as abordagens citadas acima, todas apresentam maiores dificuldades para realizar o reconhecimento facial diante de oclusões contínuas nas regiões dos olhos e boca. Tal fato deve-se a característica holística das abordagens. Na prática, o vetor de características holístico não mostra-se tão robusto a variações de iluminação, expressão, posição e deformações locais [17] [18] [19]. Buscando resolver essa dificuldade, nesse artigo é proposto a utilização do banco de filtro de Gabor curvo que extrai características locais da imagem em múltiplas frequências e orientações. Além disso, a magnitude das respostas do banco de filtro de Gabor curvo são segmentadas em blocos não sobrepostos. Em seguida é aplicada a entropia, ressaltando características locais na representação final da imagem.

II. REFERENCIAL TEÓRICO

A. Representação por Gabor

Extração de características é um tipo de redução de dimensionalidade que tem como objetivo representar partes importantes de uma imagem de forma eficiente, constituindo um vetor de características compacto [20]. Esse método é utilizado quando uma imagem é de dimensão elevada e tornase necessário sua representação como um vetor reduzido de forma a concluir rapidamente atividades como reconhecimento e seguimento.

Deteccção e extração de características bem como rastreamento são usualmente combinados para resolver problemas comuns da visão computacional, dentre eles: deteccção e reconhecimento de objetos, deteccção e reconhecimento de faces e classificação de texturas [21].

O Filtro de Gabor 2D é uma onda senoidal modulada por uma Gaussiana, expressa segundo a Equação(1) [20]:

$$\varphi_{\pi(q,\theta,\gamma,\eta)}(x,y) = \frac{q^2}{\pi\gamma\eta} \exp(-(\alpha^2 x'^2 + \beta^2 y'^2)) \exp(j2\pi qx')$$

$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta \quad (1)$$

na qual q é a frequência no plano de onda, é o ângulo, medido no sentido anti-horário, entre a direção de propagação do plano de onda e o eixo x , α é a Gaussiana no maior eixo paralelo a onda, e β é a gaussiana no menor eixo perpendicular a onda. $\alpha = q$ e $\beta = q$ são definidos para manter a proporção constante. Para extrair características de uma imagem, normalmente um conjunto de filtros de Gabor são utilizados em diversas frequências e orientações:

$$\varphi_{u,v} = \varphi_{\pi(g_u,\theta_v,\gamma,\eta)}, g_u = g_{max}/\sqrt{2^u},$$

$$\theta_v = \frac{v}{8}\pi \quad (2)$$

$$u = 0, \dots, U-1, v = 0, \dots, V-1$$

O reconhecimento facial destaca-se como uma das mais importantes aplicações do Filtro de Gabor [22]. Quando é

realizada a convolução de uma imagem facial com um conjunto de filtro de Gabor, a resultante é uma imagem que está qualificada para ser aplicada numa técnica de reconhecimento. A equação (1) possui duas partes, uma real (R) e outra imaginária (I) que podem ser definidas por:

$$\varphi(x,y) = R_{(x,y)} + I_{(x,y)} \quad (3)$$

enquanto que a magnitude é expressa pela equação:

$$|\varphi(x,y)| = |R_{(x,y)}^2 + I_{(x,y)}^2|^{1/2}. \quad (4)$$

Os filtros de gabor, tem sido usados amplamente em análise de padrões de aplicações. Dentre as mais importantes características da magnitude do filtro de Gabor, está sua invariância a iluminação, rotação, escala, e translação. Além disso, ele é pouco afetado por distúrbios de fotografia, como mudanças de iluminação e ruído de imagem.

1) *Gabor curvo*: Uma imagem de face é composta de uma conjunto de linhas curvas, ao invés de contornos retos. Na análise de imagens faciais, componentes importantes para a identificação tais como olhos, nariz, e bochechas tem formas curvas. Assim, infere-se que um filtro de Gabor curvo é mais apropriado do que o tradicional Gabor. Conforme é detalhado na Figura 2 na qual ao se sobrepor o filtro ao formato das linhas do rosto é possível deduzir que é possível extrair dados mais representativos da imagem [2].

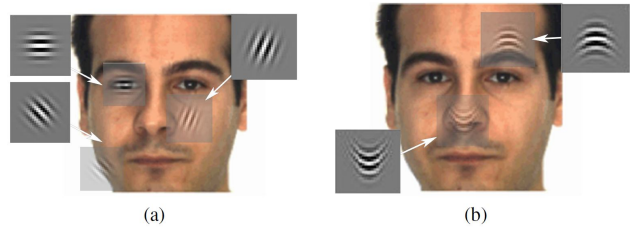


Figura 2 - Filtro de Gabor tradicional (a) e o filtro de Gabor curvo (b), aplicado nos olhos nariz e bochechas

A partir da equação (1), obtêm-se a equação (5) representando o filtro de Gabor curvo, na qual c indica o grau de curvatura:

$$\varphi_{\nu,\mu,c}(x,y) = \frac{k^2}{\sigma^2} \exp(-\frac{k_v^2}{2\sigma^2} ((x \cos \varphi_\mu + y \sin \varphi_\mu + c(-x \sin \varphi_\mu + y \cos \varphi_\mu))^2 + (-x \sin \varphi_\mu + y \cos \varphi_\mu)^2))$$

$$\times [\exp(ik_v(x \cos \varphi_\mu + y \sin \varphi_\mu + c(-x \sin \varphi_\mu + y \cos \varphi_\mu)^2) - \exp(-\frac{\sigma^2}{2}))], \quad (5)$$

em que $k_v = 2^{-(v+2)/2}\pi$ e $\varphi_\mu = \mu(\pi/8)$.

O banco de filtros de Gabor curvo, quando comparado com o convencional possui o dobro de orientações. Enquanto que conforme mencionado, utiliza-se 8 orientações para o filtro tradicional, o filtro curvo utiliza 16. Tal fato, deve-se a assimetria do banco de filtros curvo, detalhado na Figura 3, que para estender-se sobre toda as orientações tendo $\varphi_\mu = \mu(\pi/8)$ necessita do dobro de filtros.

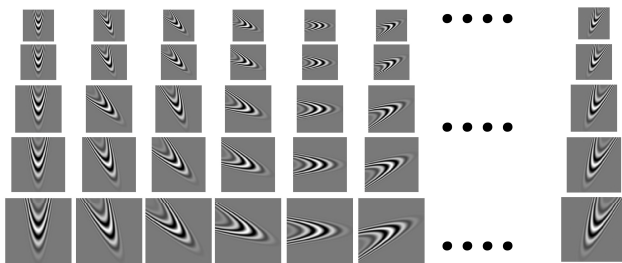


Figura 3 - Banco de filtro de Gabor curvo

Dessa forma, obtemos um novo banco de filtros de Gabor, conforme Figura 4 que contrapõe o filtro de o banco de filtro de Gabor tradicional, com sua variação de curvatura $\{c = 0,05; 0,1; 0,2\}$.

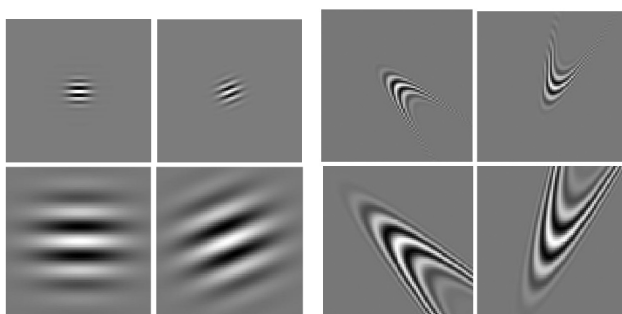


Figura 4- Banco de filtro de Gabor e banco curvo

B. Local Variance Projection Log Energy Entropy Features

Os métodos de projeção de imagem podem caracterizar a distribuição de dados ao longo de uma direção específica. Entre as diversas técnicas, a projeção de variância é uma das mais invariante a mudanças de iluminação [23] [24], propiciando uma robustez maior em ambientes não controlados. Considerando que a função discreta $f(x, y)$ é a representação da intensidade de um pixel de uma imagem na posição (x, y) e que $f(x, y)$ está definida nos intervalos $[x_1, x_2]$ e $[y_1, y_2]$, a função de projeção da variação vertical e horizontal da imagem é definida como:

$$VPF_v(x) = \frac{1}{y_2 - y_1} \sum_{y=y_1}^{y_2} [f(x, y) - IPF_v(x)]^2, \quad (6)$$

$$VPF_h(y) = \frac{1}{x_2 - x_1} \sum_{x=x_1}^{x_2} [f(x, y) - IPF_h(y)]^2, \quad (7)$$

onde $VPF_v(x)$ e $VPF_h(y)$ é a função de projeção da variância (variance projection function) sobre a vertical e a horizontal, respectivamente. $IPF_v(x)$ e $IPF_h(y)$ são a média das funções integrais na vertical e horizontal, detalhadas a seguir:

$$IPF_v(x)' = \frac{1}{y_2 - y_1} \sum_{y=y_1}^{y_2} f(x, y), \quad (8)$$

$$IPF_h(y)' = \frac{1}{x_2 - x_1} \sum_{x=x_1}^{x_2} f(x, y), \quad (9)$$

A distribuição de informação da VPF é um modelo que pode ser adaptado para extrair características de imagens de forma robusta a iluminação.

O conceito de entropia, introduzido por Shannon em [16] é inicialmente utilizado para o estudo de sistemas de comunicação. Considerando os componentes desses sistemas como elementos probabilísticos, posteriormente passa a ser utilizado em outras áreas como processamento de imagens [25]. O fundamento da entropia é

utilizar a incerteza como uma medida para descrever a informação de uma fonte. Em fonte discreta de informação com k elementos e probabilidade $p_i = \{p_1, p_2, \dots, p_k\}$ a entropia é dada por:

$$S = - \sum_{i=1}^k p_i \cdot \log p_i \quad (10)$$

Uma imagem é uma fonte de informação, tal qual um sinal unidimensional, portanto a imagem pode ser representada por meio de sua entropia. Assim, a dimensão da imagem pode ser significativamente reduzida, facilitando a extração de características e classificação.

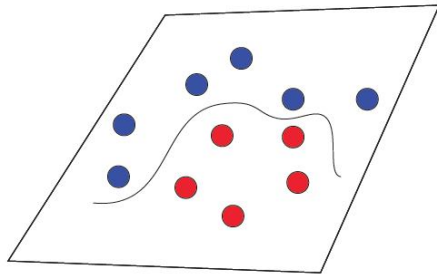
Para uma imagem $f(i,j)$ de tamanho $M \times N$, o valor da entropia de Shannon de uma imagem é definido como:

$$En = - \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{ij} \log P_{ij} \quad (11)$$

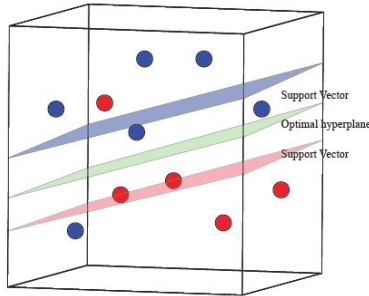
C. SVM

Máquinas de Vetores de Suporte (SVM) são uma poderosa técnica para classificação de padrões [5]. O SVM cria um hiperplano ou um conjunto de hiperplanos em um alto ou infinito espaço dimensional, que pode ser usado para classificação.

O SVM efetua o reconhecimento de padrões entre duas classes construindo um plano ou hiperplano que maximize a distância entre dois pontos próximos no conjunto dos pontos de treinamento [26], conforme Figura 5. De um conjunto de pontos $x_i \in \mathbb{R}^2$, $i = 1, 2, \dots, N$ onde cada ponto x_i pertence a uma de duas classes identificadas pelos rótulos $y_i \in \{-1, 1\}$. Considerando que os dados são linearmente separáveis, o objetivo é separar as duas classes por um hiperplano cuja distância até os vetores de suportes seja maximizada.



(a) Dados não linearmente separáveis, dispersos em 2D



(b) Dados linearmente separáveis, projetados para 3D

Figura 5 - Hiperplano gerado pelo SVM, maximizando a distância entre os

Esse hiperplano ótimo é definido por:

$$h(x) = \sum_{i=1}^l \alpha_i y_i x_i \cdot x + b, \quad (12)$$

Um novo ponto x é classificado computando a Equação 13 para uma classificação multi-classe. O sinal de d é o resultado da classificação para x , enquanto que $|d|$ é a distância de x para o hiperplano. Assim, quanto maior o valor de $|d|$, significando que o ponto é mais distante da superfície de decisão, mais confiável é o resultado da classificação.

$$d(x) = \frac{\sum_{i=1}^l \alpha_i y_i x_i \cdot x + b}{\|\sum_{i=1}^l \alpha_i y_i x_i\|} \quad (13)$$

As definições representadas pelas Equações (12) e (13) podem ser estendidas para o caso de conjunto de dados linearmente não separáveis. Dessa forma, um ponto x no espaço de entrada é mapeado para um ponto $z = (x)$ pertencente a um espaço dimensional superior no qual os dados são separáveis por um hiperplano. Uma propriedade fundamental nessa estrutura é que o produto $(x) \cdot (y)$ pode ser reescrito como uma função kernel $K(x; y)$. A superfície de decisão é estabelecida pela Equação (14):

$$m(x) = \sum_{i=1}^l \alpha_i y_i K(x, x_i) + b, \quad (14)$$

A técnica descrita para a classificação binária, pode ser empregada em duas estratégias iniciais para a resolução de problemas multi classes. Na primeira, denominada um contra todos, q SVMs são treinadas. Cada uma das SVMs separa uma única classe de todas as outras classes restantes [27] [28]. Na técnica em par, $q(q - 1)/2$ máquinas são treinadas, sendo que um par de classes é separada em cada SVM. A partir de então, vários métodos podem ser empregados para classificação multi-classes. Na implementação do SVM utilizada nesse artigo, a biblioteca libSVM [29], adota o mecanismo proposto inicialmente para treinamento em par, de redes neurais detalhado em [30] e aplicada posteriormente para SVMs em [31].

III. ABORDAGEM PROPOSTA

A abordagem visa o reconhecimento facial utilizando o banco de filtro de Gabor curvo para a extração do vetor de características, junto com a técnica de projeção local de variância log energy entropy. Ela divide-se em 5 componentes principais: aplicação do filtro de gabor curvo, segmentação em blocos, uso da entropia, treinamento e classificação com o SVM. Os componentes propostos e suas interações são vistos na Figura 6 e detalhados a seguir.

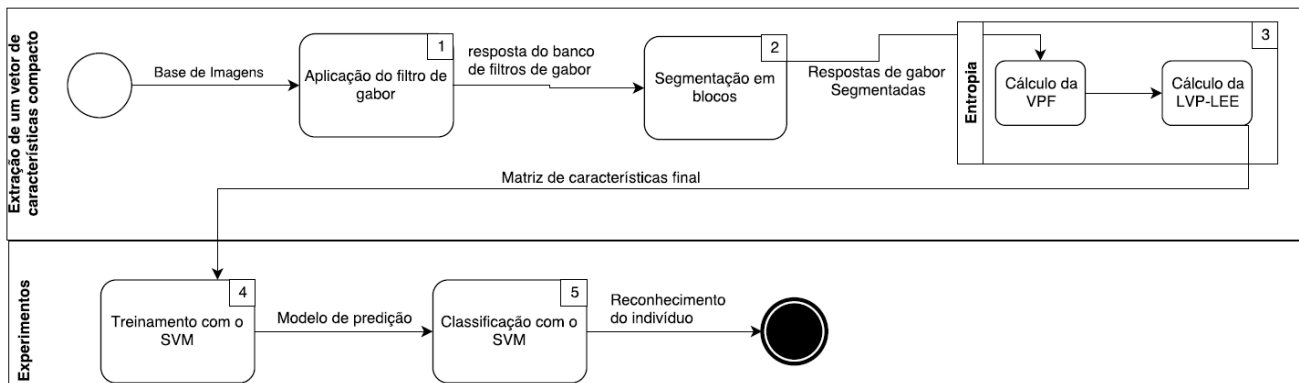


Figura 6 - Fluxograma geral da abordagem proposta

O filtro de gabor, tem sido usado amplamente em análise de padrões de aplicações [22]. Dentre as mais importantes características da magnitude do filtro de gabor, está sua invariância a iluminação, rotação, escala, e translação. Além

disso, ele é pouco afetado por distúrbios de fotografia, como mudanças de iluminação e ruído de imagem. Constitui-se assim, a etapa 1 da abordagem, sendo responsável pela

representação inicial da imagem e geração da magnitude das respostas do banco de filtro de Gabor curvo.

Na etapa 2 é efetuado um processamento com o objetivo de preparar os dados para o processamento da entropia. Considerando que imagens de diferentes indivíduos podem possuir a mesma sequência de histograma e compartilhar da mesma intensidade de entropia ou da log energy entropy, as duas imagens poderão ser identificadas incorretamente como pertencentes a mesma pessoa [32]. Uma abordagem alternativa para resolução desse problema é a segmentação da imagem da face em pequenos blocos não sobrepostos, pois todos os blocos de diferentes imagens de faces possuem distribuição diferente. Essa técnica de divisão é rápida, simples e não se baseia na detecção de importantes áreas da face como olhos, nariz e boca [33] e pode ser vista na Figura 7. A partir de então a abordagem passa a visualizar apenas os blocos individualmente ao invés da imagem toda, permitindo uma maior resistência a mudanças efetuadas em outras áreas.

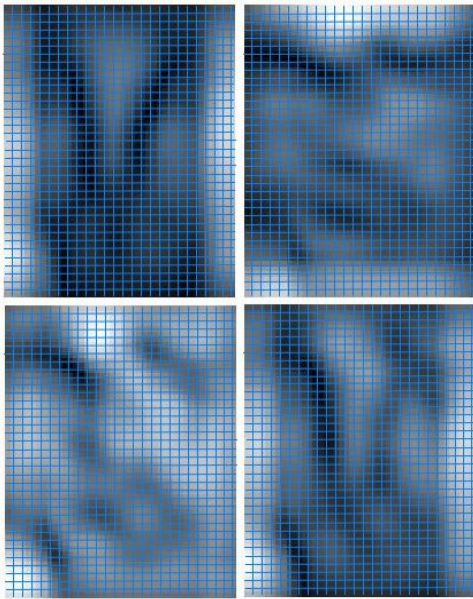


Figura 6 - Segmentação da magnitude das resposta do filtro de Gabor curvo, em blocos de 4x4 *P pixels*

Em seguida, tem-se a terceira etapa que consiste em utilizar a combinação da local variance projection log energy entropy features apresentada em [32], como redutor de dimensionalidade do vetor de características obtido a partir do filtro de Gabor.

A redução de dimensionalidade na extração de características tem dois motivos segundo [34]. O primeiro motivo é a redução do tamanho do vetor de características, propiciando uma significativa diminuição no custo computacional gerado durante a classificação. Em segundo lugar há uma maior robustez e confiabilidade por parte do classificador, pois são utilizadas no processo de identificação apenas as principais características da imagem. Em um vetor de características não processado, todos os dados possuem a mesma importância sendo tratados da mesma forma pelo

classificador. Assim, dados sem grande representatividade para a imagem, poderiam assumir um papel fundamental para o classificador. Nessa etapa 3, temos a geração do vetor de características final compacto, representação da imagem original que é a dada como entrada da abordagem.

A. Experimentos

No experimento proposto, utiliza-se o filtro de Gabor curvo, com 5 frequências e 16 orientações, bem como o filtro tradicional, com 5 frequências e 8 diferentes orientações ($v = 5$; $u = 8$), conforme trabalhos estado da arte [20] e [34]. O tamanho de cada filtro de Gabor é de 9×9 pixels. A mesma proposta é experimentada com os dois tipos de filtros para comparar os resultados obtidos.

A resposta do banco de filtros de Gabor curvo para o resultado da convolução de uma imagem com o banco de filtros de Gabor produz novas imagens. A partir dessas imagens resultantes, pode-se em sequência aplicar alguma técnica para redução de dimensionalidade. Aqui, optou-se pela Local Variance Projection Log Energy Entropy (LVP-LEE) [23] [24]. Para cada uma das imagens resultantes, aplicamos o método baseado na entropia, duplicando dessa forma o número de novas imagens. Ao ser aplicada a LVP-LEE em cada resposta do banco de filtros de Gabor, duas novas imagens são produzidas representando a entropia da projeção da variância sobre a horizontal e vertical.

Considerando uma imagem $I(x, y)$ com dimensões $M \times N$, a imagem da face é primeiramente dividida em L pequenos blocos não sobrepostos. O tamanho de cada bloco representado por $f''(x, y)$ é dado por $M/\sqrt{L} \times N/\sqrt{L}$. Baseando-se nesse princípio e tomando o intervalo $[k_1; k_2]$ como as posições dos bloco sobre a função f'' , pode-se aplicar a técnica da projeção da variância em cada um dos blocos pertencentes as linhas (U) e colunas (V) definidas da seguinte forma:

$$U[i] = \frac{1}{k_2 - k_1} \sum_{k=k_1}^{k_2} [f''(i, k) - g_v(i)]^2, 1 \leq i \leq M/\sqrt{L} \quad (15)$$

onde:

$$g_v(i) = \frac{1}{y_2 - y_1} \sum_{y=y_1}^{y_2} f''(i, y), \quad (16)$$

Em seguida, é calculada a log energy entropy para cada uma das linhas e colunas resultantes da Equação 15:

$$H_{RP} = - \sum_{i=1}^{M/\sqrt{L}} e_i \log(e_i) \quad (17)$$

onde

$$e_i = U[i] / \sum_{i=1}^{M/\sqrt{L}} U[i] \quad (18)$$

Para uma única imagem qualquer, são produzidos ao final $2L$ valores, representados na forma de duas imagens conforme a Figura 8. Essas características são nomeadas como *vertical*

variance projection entropy vector e *horizontal variance projection entropy vector* e após serem concatenadas em um só vetor estão aptas a serem utilizadas por um classificador de forma a reconhecer uma face.

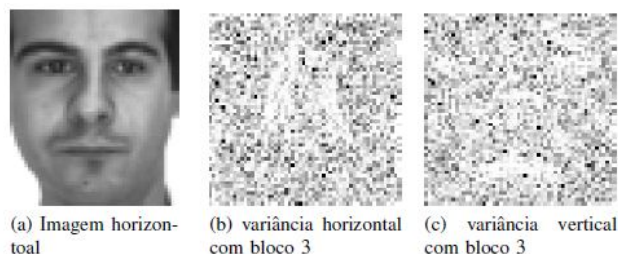


Figura 7- Imagem frontal extraída da AR Face, com suas respectivas

Como ferramenta de apoio a execução do SVM como classificador, utilizou-se a biblioteca LIBSVM [29], com o kernel RBF (*Radial basis function*). Desse ponto, inicia-se o processo de treinamento e classificação das imagens com o SVM, conforme exposto nas etapas 4 e 5 da Figura 6.

IV. RESULTADOS E DISCUSSÕES

Esta seção apresenta detalhes da aplicação das abordagens na AR Face Database [35] e os resultados obtidos. Constituída de mais de 4000 imagens coloridas de um total de 166 pessoas, a AR Face Database é uma importante base para verificação de acurácia em algoritmos de reconhecimento facial. Os experimentos de reconhecimento facial foram avaliados por meio da utilização da técnica do 10-fold cross-validation para a medição da acurácia no reconhecimento dos indivíduos definidos na base de imagens. Dentre as diversas técnicas de reconhecimento facial, encontram-se as abordagens SRC [6], LRC [7], CESR [8], FSS [9], SGLasso [10] e DCC [11] que representam o estado da arte e foram utilizados para comparar com as duas abordagens propostas neste trabalho: Gabor com entropia e Gabor Curvo com entropia. Cada um desses algoritmos possui características próprias que merecem ser detalhadas.

A. AR Face

As imagens dessa base, correspondem a fotos de faces frontais com diferentes expressões faciais, condições de iluminação e oclusões compreendendo 70 homens e 56 mulheres. A principal característica dessa base de faces é a oclusão gerada pelo uso de óculos e cachecol por parte dos indivíduos representados. Para a realização dos nossos experimentos, as imagens foram redimensionadas para 48x56 pixels, de forma a aproximar-se das dimensões utilizadas pelos demais algoritmos que são usados em comparação. A Figura 9 representa os detalhes da base: Figura 9a, expressão neutra, Figura 9b sorriso, Figura 9c raiva, Figura 9d susto, Figura 9e luz a esquerda, Figura 9f luz a direita e Figura 9g luz frontal. A AR Face foi segmentada em 5 diferentes subconjuntos, de forma a agrupar características distintas nos testes efetuados. Imagens de faces com diferentes variações de expressão facial,

iluminação e oclusão são agrupadas possibilitando uma experimentação mais robusta da abordagem proposta em diversos cenários. A seguir cada subconjunto é detalhado:

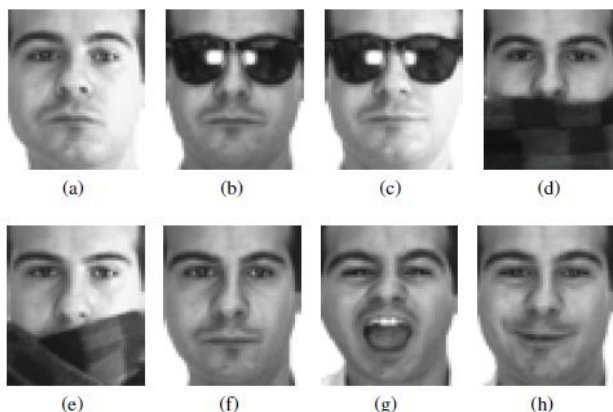


Figura 9 - Faces extraídas da Ar Face Database

1) *Varição apenas de iluminação*: a imagem de uma face sofre efeitos bastante complexos a partir de variações de iluminação, dificultando a etapa de reconhecimento facial. Mudanças no sentido da iluminação causam alterações na localização e forma das sombras, bem como variações nos realces e contrastes. O subconjunto das imagens de 5 a 7 e 18 a 20 (Figura 9a) na Ar Face referentes a cada indivíduo compreendem apenas essa variação de iluminação. A Tabela 1 contém os valores referentes a acurácia das abordagens no reconhecimento facial dos indivíduos da AR Face. Os melhores resultados nos termos de acurácia são os obtidos por CC e DCC (80% e 71% respectivamente). Eles representam um significativo avanço diante das demais abordagens, que alcançam o máximo de acerto de 48,74%. As duas abordagens propostas, sobrepõem os outros métodos por percentual significativo de até 12,5%, obtendo 82,17% de acerto com o uso do Gabor com a entropia e alcançando 84,33% com o Gabor Curvo com a entropia. Verifica-se assim que a invariância do filtro de Gabor a mudanças na iluminação é de fundamental importância para o resultado obtido nas duas abordagens.

2) *Oclusão por óculos de sol*: um dos maiores desafios no reconhecimento facial é a oclusão parcial da face. Em um ambiente não controlado, a face de uma pessoa pode ser facilmente obstruída por acessórios como óculos, cachecol ou chapéu ou por obstáculos externos. O subconjunto formado pelas imagens de número 8 e 21 (Figura 9b), constituem uma oclusão frontal causada por um óculos de sol, possuindo cerca de 20% da área total obstruída. De acordo com a Tabela I, observamos que DC e DCC apresentam os melhores percentuais, com 74,79% e 72,69% de acerto. A proposta de uso do Gabor e entropia não mostrou-se eficiente nesse teste.

Tabela 1- Comparação dos resultados da abordagem proposta, com abordagens do estado da arte

Abordagem	LRC	SRC	CESR	FSS	SGLasso	DC	CC	DCC	Gabor Curvo + Entropia	Gabor + Entropia
Iluminação	31.37	45.80	48.74	44.96	38.24	45.24	79.83	71.01	84.33	82,17
Oclusão por óculos de sol	25.21	28.99	68.49	28.99	21.85	74.79	3.78	72.69	50.00	0
Iluminação + óculos de sol	8.19	15.13	20.80	14.50	11.55	23.95	3.15	22.48	73.75	65,67
Oclusão por cachecol	94.96	95.38	96.64	95.38	93.28	97.06	68.49	97.06	86	83,50
Iluminação + Oclusão por cachecol	18.28	29.41	36.76	27.31	21.22	29.20	63.87	45.80	89,25	85,50
Mudança de expressão facial	x	x	x	x	x	x	x	x	91.5	91,75
Base completa	x	x	x	x	x	x	x	x	98.15	97,88

Tal fato deve-se a característica do classificador utilizado, o SVM, que necessita de um número maior de amostras por classes para que possa classificar corretamente os indivíduos, mostrando que 2 imagens são insuficientes para essa abordagem. Para compatibilizar os testes com os já realizados na literatura, adotou-se o mesmo particionamento da base de imagens, segregando as oclusões por cachecol, óculos de sol e óculos de sol com variação de iluminação, assim restringiu-se a 2 imagens por cada indivíduo nesse experimento. Contudo, ao fazer o uso do filtro de Gabor curvo com a entropia obtém-se um resultado de 50%, evidenciando um avanço na extração das características mais relevantes da face, pois o crescimento na taxa de acerto ocorreu sem quaisquer mudanças na etapa da classificação, tendo apenas a melhor representação da face como mudança.

3) *Iluminação + óculos de sol*: esse experimento combina o desafio gerado pela variação da iluminação presente no primeiro teste, com a oclusão do segundo. A Figura 9c contém exemplos extraídos das imagens 9, 10, 22 e 23 que demonstram um acréscimo no grau de dificuldade no reconhecimento desse subconjunto, confirmado pelas taxas de acerto obtidas pelas abordagens analisadas. Os melhores resultados encontrados 23,95%(DC) e 22,48% (DCC) são bastante inferiores aos testes que apresentam apenas a variação de iluminação ou oclusão por óculos de sol, contudo a abordagem proposta com Gabor e Entropia gera um acerto de 65,67% seguido por 73,75% do Gabor curvo e entropia, evidenciando a evolução com relação as técnicas anteriormente propostas, quando há uma combinação de intensa variação de luz e oclusão.

4) *Oclusão por cachecol*: com cerca de 45% de oclusão da face gerada pelo uso de um cachecol, esse subconjunto é composto pelas imagens 11 e 24 (Figura 9d) da ARFace. A oclusão criada ocorre em toda a parte inferior da imagem assim considera-se que a imagem pode ser cortada em 50% da sua altura, obtendo uma nova imagem que será utilizada na classificação. Nesse experimento, todos os métodos empregados obtiveram taxas de acerto elevados, sendo que 6 deles apresentaram valores superiores a 94%. Nossas abordagens, no entanto produziram valores de 83,50% para Gabor e entropia e 86% para Gabor Curvo e entropia. Tais dados justificam-se pela significativa redução no tamanho da imagem final utilizada nos testes. Sugerindo que o nível de redução de dimensionalidade aplicado às imagens deve ser ajustado com relação ao tamanho da imagem.

5) *Iluminação + Oclusão por cachecol*: adicionando-se a variação de iluminação à oclusão por cachecol nas imagens 12, 13, 25 e 26 (Figura 9e) , propiciando um cenário mais complexo para o reconhecimento facial. Para esses caso o Gabor com entropia alcança 85,50% de acerto superado pelo Gabor curvo e entropia com 89,25%, taxas acima dos 63,87% alcançados por DC.

6) *Mudança de expressão Facial*: outro fator fundamental no reconhecimento de imagens é a robustez às mudanças na expressão facial. Tomando como testes o subconjunto formado pelas imagens de 1 a 4 e 14 a 17 (Figura 9f), segregamos as faces que representam expressões neutras, sorriso, raiva e susto, possibilitando a verificação em um ambiente mais próximo do mundo real. Os valores obtidos com Gabor Curvo e entropia destacam-se por alcançar 91,25% de acerto.

7) *Base completa*: por fim, faz-se necessário a avaliação das duas abordagens propostas com todas as imagens da AR Face (Figura 9), a fim de conseguir a simulação de um ambiente não controlado, com as possíveis ocorrências encontradas no mundo real. Nesse contexto obtivemos 92,38% de acerto para Gabor e entropia e 98,15% para Gabor Curvo e entropia com a abordagem proposta.

B. Discussões

Os resultados obtidos pelas abordagens propostas foram superiores em todos os experimentos que possuíam variação direta de iluminação, ressaltando assim a sua robustez a esse tipo de ruído. Vale ressaltar que nos testes os principais problemas decorrem do baixo número de imagens por indivíduo como no caso da oclusão por óculos de sol, na qual há apenas 2 imagens por pessoa ou pela reduzida dimensão das imagens como na oclusão por cachecol na qual a imagem é reduzida em 50% pela oclusão. Quanto aos 2 últimos testes, com mudança de expressão facial e com todas as imagens da ARface, obtém-se um contexto mais próximo ao encontrado no mundo real, com diversas variações para um mesmo rosto. Nesse aspecto, alcançou-se resultados superiores a 98% de acerto, caracterizando a robustez do uso do Gabor curvo com a entropia, conforme proposto nesse trabalho.

Observa-se em todos os testes apresentados que a adição do filtro de Gabor curvo em relação ao filtro tradicional propicia um avanço na acurácia encontrada. Essa diferença entre as duas abordagens pode assumir valores sutis, como no experimento 1 com variação de 2,16% podendo chegar até a 50% como no experimento 2. Tal fato justifica-se pela adição de mais

características relevantes na representação da face, que são evidenciadas em situações com oclusão.

Os testes foram efetuados com a AR Face, base muito empregada na literatura para testes de acurácia das abordagens de reconhecimento facial. Apesar de ser uma base criada em um ambiente controlado, ela adquire características de um ambiente não controlado, incluindo variação de iluminação, oclusão e expressões faciais, consolidando-se como uma base de referência. Os experimentos efetuados sobre a AR Face nas duas propostas apresentadas nesse trabalho, possuem como etapa que requer maior poder computacional a aplicação do filtro de Gabor. Nesse caso, a utilização do filtro de Gabor curvo acentua esse aspecto, pois ele possui 4 variações de grau de curvatura. Contudo, os ganhos obtidos pela inclusão de novos filtros curvos, superam o crescimento no custo de execução da solução.

V. CONCLUSÕES

Nesse artigo, foi apresentada uma abordagem para realização de reconhecimento facial, combinando um banco de filtro de Gabor curvo e a entropia. Os índices de acerto superiores, na maior parte dos casos de teste, aos algoritmos no estado da arte comprovam a eficiência da utilização do Filtro de Gabor curvo em conjunto com a entropia. Além disso, obteve-se um vetor de características compacto (devido ao uso da entropia como redutor de dimensionalidade) e robusto a variações de iluminação (caracterizado pela aplicação do filtro Gabor curvo) e a oclusões (decorrente da segmentação em blocos).

Os resultados obtidos referem-se aos experimentos nos quais as imagens estão posicionadas de forma frontal, variando iluminação e com oclusão. Segundo os resultados, na redução do vetor de características obteve-se um ganho considerável na acurácia de reconhecimento facial, devido a otimização da representação da imagem. Dentre as particularidades apresentadas pela abordagem proposta, vale ressaltar que a abordagem obtém melhores resultados quando aplicada a bases faciais que contêm um número significativo de imagens por indivíduo, com pouca variação entre elas. Essa limitação apresentada, pode ser reduzida em trabalhos futuros, aplicando-se outros tipos de classificadores. O principal aspecto a ser buscado em outro algoritmo de classificação é uma maior flexibilidade a amostras que possuam uma diferença expressiva dentro da mesma classe.

REFERENCIAS

- [1] U. Srinivas, Y. Suo, M. Dao, V. Monga, and T. D. Tran, "Structured sparse priors for image classification," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1763–1776, 2015.
- [2] W. Hwang, X. Huang, S. Z. Li, and J. Kim, "Face recognition using Extended Curvature Gabor classifier bunch." *Pattern Recognition*, vol. 48, no. 4, pp. 1247–1260, 2015. [Online]. Available: <http://dblp.uni-trier.de/db/journals/pr/pr48.html#HwangHLK15>
- [3] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
- [4] R. Chellappa, J. Ni, and V. M. Patel, "Remote identification of faces: Problems, prospects, and progress." *Pattern Recognition Letters*, vol. 33, no. 14, pp. 1849–1859, 2012.
- [5] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1995.
- [6] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, Feb 2009.
- [7] I. Naseem, R. Togneri, and M. Bennamoun, "Linear Regression for Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2106–2112, Nov 2010.
- [8] R. He, W. S. Zheng, and B. G. Hu, "Maximum Correntropy Criterion for Robust Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1561–1576, Aug 2011.
- [9] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Advances in neural information processing systems*, 2006, pp. 801–808.
- [10] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.
- [11] Z. R. Lai, D. Q. Dai, C. X. Ren, and K. K. Huang, "Discriminative and Compact Coding for Robust Face Recognition," *IEEE Transactions on Cybernetics*, vol. 45, no. 9, pp. 1900–1912, Sept 2015.
- [12] A. Kar, D. Bhattacharjee, D. K. Basu, M. Nasipuri, and M. Kundu, "Human face recognition using gabor based kernel entropy component analysis," *International Journal of Computer Vision and Image Processing (IJCVIP)*, vol. 2, no. 3, pp. 1–20, 2012.
- [13] R. Bellman, *Adaptive Control Processes*. Princeton University Press, 1961.
- [14] L. Valiant, "A Theory of the Learnable," *Communications of the ACM*, vol. 27, no. 11, pp. 1134–1142, 1984.
- [15] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press, 1995.
- [16] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, The, vol. 27, no. 3, pp. 379–423, July 1948.
- [17] Y. Wong, M. T. Harandi, C. Sanderson, and B. C. Lovell, "On robust biometric identity verification via sparse encoding of faces: Holistic vs local approaches," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–8.
- [18] D.W. Piepers and R. A. Robbins, "A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature," *Frontiers in psychology*, vol. 3, 2012.
- [19] R. Weng, J. Lu, and Y. P. Tan, "Robust point set matching for partial face recognition," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1163–1176, March 2016.
- [20] L. Shen, L. Bai, and M. C. Fairhurst, "Gabor wavelets and General Discriminant Analysis for face identification and verification." *Image Vision Comput.*, vol. 25, no. 5, pp. 553–563, 2007.
- [21] S. Meshgini, A. Aghagolzadeh, and H. Seyedarabi, "Face recognition using Gabor-based direct linear discriminant analysis and support vector machine." *Computers & Electrical Engineering*, vol. 39, no. 3, pp. 727–745, 2013.
- [22] C. Liu and H. Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition," *Trans. Img. Proc.*, vol. 11, no. 4, pp. 467–476, Apr. 2002.

- [23] Z.-H. Zhou and X. Geng, "Projection functions for eye detection," *Pattern recognition*, vol. 37, no. 5, pp. 1049–1056, 2004.
- [24] G.-C. Feng and P. C. Yuen, "Variance projection function and its application to eye detection for human face recognition," *Pattern Recognition Letters*, vol. 19, no. 9, pp. 899–906, 1998.
- [25] C.-I. Chang, Y. Du, J. Wang, S.-M. Guo, and P. Thouin, "Survey and comparative analysis of entropy and relative entropy thresholding techniques," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 153, no. 6, pp. 837–850, Dec 2006.
- [26] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Eighth IEEE International Conference on Computer Vision*, 2001. *ICCV 2001*, vol. 2. IEEE, 2001, pp. 688–694.
- [27] B. Schölkopf, C. Burges, and V. Vapnik, "Extracting support data for a given task," in *First International Conference on Knowledge Discovery and Data Mining*. Menlo Park: AAAI Press, 1995.
- [28] C. Cortes and V. Vapnik, "Support Vector Networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [29] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.
- [30] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-Layer Learning Revisited: A Stepwise Procedure for Building and Training a Neural Network," in *Neurocomputing: Algorithms, Architectures and Applications*, ser. NATO ASI Series, F. Fogelman Soulié and J. Héroult, Eds. Springer-Verlag, 1990, vol. F68, pp. 41–50.
- [31] U. H.-G. Kreßel, "Pairwise classification and support vector machines," in *Advances in kernel methods*. MIT Press, 1999, pp. 255–268.
- [32] J.-S. Zhang and C.-J. Chen, "Local variance projection log energy entropy features for illumination robust face recognition," in *International Symposium on Biometrics and Security Technologies*, 2008. *ISBAST 2008*. IEEE, 2008, pp. 1–5.
- [33] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "Pcanet: A simple deep learning baseline for image classification?" *Image Processing, IEEE Transactions on*, vol. 24, no. 12, pp. 5017–5032, 2015.
- [34] M. B. A. Haghighat and E. Namjoo, "Evaluating the informativity of features in dimensionality reduction methods," in *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*. IEEE, 2011, pp. 1–5.
- [35] A. M. Martinez, "The AR face database," *CVC Technical Report*, vol. 24, 1998.