



Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Entropia de projeção do Gabor Curvo com *Random Forest* e SVM para reconhecimento de face

Eucassio Gonçalves Lima Júnior

Número de Ordem PPGCC: M042

Teresina-PI, Junho de 2017

Eucassio Gonçalves Lima Júnior

Entropia de projeção do Gabor Curvo com *Random Forest* e SVM para reconhecimento de face

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Ricardo de Andrade Lira Rabêlo

Coorientador: Cornélia Janayna Pereira Passarinho

Teresina-PI

Junho de 2017

Eucassio Gonçalves Lima Júnior

Entropia de projeção do Gabor Curvo com *Random Forest* e SVM para reconhecimento de face/ Eucassio Gonçalves Lima Júnior. – Teresina-PI, Junho de 2017-

51 p. : il. (algumas color.) ; 30 cm.

Orientador: Ricardo de Andrade Lira Rabêlo

Dissertação (Mestrado) – Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação, Junho de 2017.

1. Filtro de Gabor curvo. 2. Reconhecimento facial. I. Ricardo de Andrade Lira Rabêlo II. Cornelia Janayna Pereira Passarinho III. Universidade Federal do Piauí – UFPI IV. Entropia de projeção do Gabor Curvo com *Random Forest* e SVM para reconhecimento de face.

CDU 02:141:005.7

Eucassio Gonçalves Lima Júnior

Entropia de projeção do Gabor Curvo com *Random Forest* e SVM para reconhecimento de face

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Trabalho aprovado. Teresina-PI, 09 de Junho de 2017:

Ricardo de Andrade Lira Rabêlo
Orientador

Cornélia Janayna Pereira Passarinho
Co-Orientadora

Rodrigo de Melo Souza Veras
Convidado 1

Evandro Ottoni Teatini Sales
Convidado 2

Teresina-PI
Junho de 2017

*Aos meus pais Eucassio Gonçalves e Rosa Maria e a minha irmã Rosane,
por sempre estarem comigo em todos os momentos.*

Agradecimentos

Agradeço a Deus.

Aos meus pais, a quem devo grande parte de minhas conquistas e a minha irmã, sempre presente.

Agradeço aos meus orientadores, Ricardo de Andrade Lira Rabêlo e Cornelia Janayna Pereira Passarinho, por todos os conselhos, pela paciência e ajuda nesse período.

Ao companheiro de pesquisa Luis Henrique Silva Vogado pelas contribuições no trabalho.

*“Saber muito não lhe torna inteligente.
A inteligência se traduz na forma que você recolhe, julga, maneja e, sobretudo,
onde e como aplica esta informação.”
(Carl Sagan)*

Resumo

O reconhecimento facial tem por objetivo primordial identificar uma pessoa baseada em imagens de sua face. Apesar de vários avanços obtidos, alguns desafios permanecem como problemas não resolvidos, resultantes principalmente de condições não controladas do ambiente, tais como mudanças de iluminação, oclusão, variações de expressões faciais e posição da cabeça. Este trabalho apresenta uma proposta de representação de imagens faciais para reconhecimento. A representação facial proposta é constituída de seis etapas e baseia-se na combinação do filtro de Gabor curvo, entropia, *Random Forest* e *Support Vector Machine* (SVM). Um conjunto de experimentos foi realizado, para avaliar a proposta apresentada sobre características de cenários encontrados em um ambiente real, fazendo uso das bases de imagens *AR Face*, FERET e YALE. A representação da face foi avaliada em diversos cenários considerando 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, apresentam melhorias incrementais significativas diante de abordagens disponíveis na literatura, obtendo 98,03% de acurácia para a *AR Face* completa, 97,26 para a FERET e 81,66% com a Yale com 50% de oclusão.

Palavras-chaves: entropia, filtro de gabor, oclusão, reconhecimento facial, *randomforest*, redução de dimensionalidade

Abstract

Face recognition has as its primary objective to identify a person in pictures of his face. Despite several advances, some challenges remain as unresolved problems, resulting mainly from uncontrolled environmental conditions, such as changes in lighting, occlusion, variations of facial expressions and head position. This paper presents a proposal for the representation of facial images for facial recognition, characterized by robustness to occlusion, illumination variations and facial expressions. The proposed facial representation consists of six steps and is based on the combination of the curved Gabor filter, entropy, Random Forest and Support Vector Machine(SVM). A set of experiments was carried out to evaluate the presented proposal on the characteristics of scenarios found in a real environment, making use of the data bases, FERET and YALE. The face representation was evaluated in several scenarios considering light variations, occlusion by sunglasses, variation of lighting with occlusion by sunglasses, occlusion by scarf, variation of lighting with occlusion by scarf, change of facial expression and all scenarios together. The results obtained through the experiments, present significant incremental improvements in relation to available approaches in the literature, obtaining 98.03% accuracy for full AR Face , 97.26% for FERET and 81.66% with Yale with 50% of occlusion.

Keywords: gabor filter. entropy. dimensionality reduction. occlusion. facial recognition.

Lista de ilustrações

Figura 1 – Fluxo geral de um sistema de reconhecimento facial, fonte: (LI; JAIN, 2011).	1
Figura 2 – Fluxograma geral da abordagem proposta.	4
Figura 3 – Exemplos de banco de filtros de Gabor.	11
Figura 4 – Exemplo da magnitude das respostas do Gabor.	12
Figura 5 – Filtro de Gabor clássico (a) e o filtro de Gabor curvo (b), aplicado nos olhos nariz e bochechas.	12
Figura 6 – Banco de filtro de Gabor curvo.	13
Figura 7 – Banco de filtro de Gabor e banco curvo.	14
Figura 8 – Hiperplano gerado pelo SVM, maximizando a distância entre os dados de treinamento.	16
Figura 9 – Fluxograma geral da abordagem proposta.	20
Figura 10 – Cada bloco da resposta da magnitude do Gabor curvo tem a abordagem baseada em entropia calculada, resultando em duas novas representações (Etapa 3).	23
Figura 11 – Seleção de características com <i>random forest</i> (Etapa 4).	26
Figura 12 – Variação da acurácia pelo número de árvores para a FERET.	27
Figura 13 – Faces extraídas da <i>Ar Face Database</i>	30
Figura 14 – Comparação de diversas abordagens com o subconjunto da <i>AR Face</i> variação de iluminação.	32
Figura 15 – Comparação de diversas abordagens com o subconjunto da <i>AR Face</i> com oclusão por óculos de sol.	34
Figura 16 – Comparação de diversas abordagens com o subconjunto da <i>AR Face</i> com variação de iluminação e oclusão por óculos de sol.	35
Figura 17 – Comparação de diversas abordagens com o subconjunto da <i>AR Face</i> com oclusão por cachecol.	37
Figura 18 – Comparação de diversas abordagens com o subconjunto da <i>AR Face</i> com variação de iluminação e oclusão por cachecol.	38
Figura 19 – Faces extraídas da <i>FERET</i>	40
Figura 20 – Comparação de diversas abordagens com a FERET.	41
Figura 21 – Imagem de um babuíno utilizada para gerar oclusão.	42
Figura 22 – Faces extraídas da Yale com oclusão.	42

Lista de tabelas

Tabela 1 – Busca pelos melhores valores de s e k	28
Tabela 2 – Comparação dos resultados da EPGC, com abordagens da literatura. . .	33
Tabela 3 – Precisão e <i>recall</i> da EPGC para os subconjuntos da <i>AR Face</i>	36
Tabela 4 – Comparação da acurácia de diferentes algoritmos sobre a FERET. . . .	39
Tabela 5 – Comparação dos resultados da EPGC, com abordagens sobre a base Yale com oclusão.	41
Tabela 6 – Precisão e <i>recall</i> referentes a Yale.	43
Tabela 7 – Comparativo de abordagens. Na primeira não há redução do vetor de características. Na segunda e terceira abordagens é utilizado o <i>Random Forest</i> , variando o número de vezes que a Etapa 3 da EPGC é executada. Por fim, na quarta abordagem utiliza-se o <i>Extremely Randomized Trees (ERT)</i>	44

Lista de abreviaturas e siglas

CESR	<i>Correntropy-based sparse representation</i>
CUDA	<i>Compute Unified Device Architecture</i>
CPU	Central de Processamento Unitário
EPGC	Entropia de projeção do Gabor Curvo
ERT	<i>Extremely Randomized Trees</i>
GPU	<i>Graphics Processing Unit</i>
DCC	<i>Discriminative and compact coding</i>
FSS	<i>Feature-sign search</i>
LEE	<i>Log Energy Entropy</i>
LRC	<i>Linear Regression Classification</i>
LVP	<i>Local Variance Projection</i>
RF	<i>Random Forest</i>
SBAI	Simpósio Brasileiro de Automação Inteligente
SGLasso	<i>Structured graphical lasso</i>
SRC	<i>Sparse Representation-based Classification</i>
SVM	<i>Support Vector Machines</i>
VPF	<i>Variance Projection function</i>

Lista de símbolos

q	Frequência no plano da onda
θ	Ângulo do filtro de Gabor
ψ	Banco de filtro de Gabor
φ	Magnitude das respostas do Gabor
ϕ	Fase das respostas do Gabor
c	Grau de curvatura do filtro de Gabor

Sumário

	Introdução	1
1	REFERENCIAL TEÓRICO	9
1.1	Representação por Gabor	9
1.1.1	Gabor curvo	11
1.2	<i>Random Forest</i>	13
1.3	<i>Local Variance Projection Log Energy Entropy Features</i>	14
1.3.1	Função de projeção de variância	14
1.3.2	Entropia	15
1.4	SVM (Máquinas de Vetores de Suporte)	15
1.5	Considerações Finais	17
2	ABORDAGEM PROPOSTA	19
2.1	Componentes da Abordagem	19
2.1.1	Aplicação do banco de filtros de Gabor curvo	19
2.1.2	Segmentação em blocos	21
2.1.3	Entropia	22
2.1.4	Seleção de características por <i>Random Forest</i>	23
2.1.5	Treinamento e classificação com o SVM	24
2.2	Considerações Finais	25
3	RESULTADOS E DISCUSSÕES	29
3.1	AR Face	29
3.1.1	Variação apenas de iluminação	30
3.1.2	Oclusão por óculos de sol	31
3.1.3	Iluminação + óculos de sol	31
3.1.4	Oclusão por cachecol	36
3.1.5	Iluminação + Oclusão por cachecol	36
3.1.6	Mudança de expressão Facial	36
3.1.7	Base completa	39
3.2	FERET	39
3.3	Yale com oclusão	39
3.4	Discussões	42
3.5	Considerações Finais	44
4	CONCLUSÕES E CONTINUIDADE DA PESQUISA	45

4.1	Continuidade da pesquisa	45
	REFERÊNCIAS	47

Introdução

O reconhecimento facial é uma ramo de pesquisa que desenvolveu-se rapidamente (SRINIVAS et al., 2015), destacando-se dentro da área de processamento e análise de imagens. Dentre as diversas aplicações que lidam com reconhecimento facial, pode-se citar: identificação de pessoas para acesso a dados restritos, controle de acesso e segurança, monitoramento de multidões, localização de fichas criminais, interatividade controlada por expressões faciais, dentre outros (LI; JAIN, 2011).

O reconhecimento facial tem por objetivo primordial identificar uma pessoa baseada em imagens de sua face. Para isso, é necessário comparar as imagens iniciais que devem ser identificadas com várias amostras de um banco de dados facial. Atualmente existem diversas fontes de imagens de face, já armazenadas para os mais variados fins, tais como: bancos de dados de fotografias de passaportes, carteiras de motoristas, documentos de identidade, redes sociais dentre outros relatados em (PHILLIPS et al., 2000), com milhões de indivíduos diferentes. Apesar dos avanços ocorridos nas últimas décadas, fruto de intensas pesquisas, alguns desafios permanecem como problemas não resolvidos (CHELLAPPA; NI; PATEL, 2012). Esses desafios não resolvidos são resultantes principalmente de condições não controladas do ambiente, tais como oclusão, posicionamento, variações de expressões faciais e mudanças de iluminação.

Definição do Problema

Um sistema de reconhecimento facial consiste em efetuar uma busca do tipo um-para-muitos comparando a imagem da face com múltiplas faces em uma base de dados (LI; JAIN, 2011). Constitui-se dessa forma, em um problema de reconhecimento de padrões visuais. O sistema de reconhecimento facial é composto de quatro módulos principais: localização e marcação da face, normalização das faces, extração de características, e combinação de característica, conforme Figura 1.

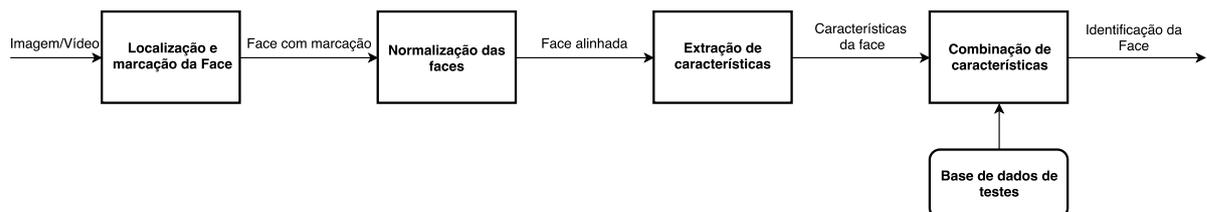


Figura 1 – Fluxo geral de um sistema de reconhecimento facial, fonte: (LI; JAIN, 2011).

O primeiro módulo, **localização e marcação da face** é responsável por receber um vídeo ou uma imagem e segmentar a face do fundo da imagem. Ao realizar essa segmentação

obtém-se a localização da face na imagem, isolando-a do fundo da imagem, descartando as demais informações do fundo, não relevantes ao problema. Diversas indicações podem ser utilizadas nesse ponto como: cor da pele, movimento (para vídeos), forma da face, aparência da face ou uma combinação desses elementos.

Em seguida é efetuada a **normalização da face**, processo no qual a face já detectada é normalizada quanto ao tamanho e postura, empregando transformações geométricas. Assim, a imagem da face é rotacionada e cortada de forma a obter-se uma nova imagem padronizada quanto às demais do banco de dados. Igualmente, as propriedades fotométricas da face, como iluminação e nível de cinza, são normalizadas.

No módulo três, **extração de características**, de posse da face normalizada em uma imagem padronizada, deve-se extrair as características faciais mais relevantes, de forma que possam distinguir indivíduos diferentes. Tais características devem ter o mínimo de variação a transformações geométricas e fotométricas e formam o vetor de características, uma representação computacional da face, otimizada para reconhecimento facial.

Por fim é efetuada a **combinação de características**. Nesse módulo, o vetor de características, extraído na etapa anterior, é comparado com os vetores das demais imagens e tem como saída, a identificação ou não do indivíduo. A definição da medida de similaridade é um dos grandes desafios em aberto. Como proposta de solução são empregadas diversas técnicas de classificação de dados.

Segundo (LI; JAIN, 2011), em um ambiente ideal, uma representação das características faciais de uma pessoa deve possuir os seguintes requisitos: 1) diferenciar bem indivíduos distintos (grande variação inter-classes), enquanto que aceita variações entre as representações de um mesmo indivíduo (variações intra-classes); 2) permitir uma extração com baixo custo computacional a partir de imagens da face, possibilitando um processamento rápido; e 3) possuir um baixo espaço dimensional, propiciando uma redução no custo computacional com o classificador.

Representar características faciais não é uma tarefa trivial, tornando-se um dos principais problemas na área do reconhecimento facial (CHEN et al., 2015). Existem variáveis não controladas que influenciam diretamente na obtenção dessa representação, como: escala, orientação, posição, expressões faciais, condições de iluminação, idade e o uso de acessórios (óculos, cachecol, chapéu, dentre outros).

Um aspecto de um sistema robusto de reconhecimento facial é a representação das características de uma imagem por meio de um vetor com baixa dimensionalidade. Métricas como armazenamento, tempo de recuperação e identificação devem ser otimizadas nesse tipo de sistema. A técnica de representação geralmente inicia-se com um procedimento de redução de dimensionalidade, pois a alta dimensionalidade do vetor que representa uma

imagem original, dificulta ou mesmo inviabiliza o classificador de obter bons resultados (KAR et al., 2012).

Conforme proposto inicialmente em (BELLMAN, 1961), a medida que aumentasse o número de características que devem ser consideradas, cresce exponencialmente o número de características do conjunto de dados. Para um algoritmo de aprendizagem de máquina (ALPAYDIN, 2016), definir um modelo preciso de predição com diversas características torna-se mais complexo (WITTEN et al., 2016). 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 características (VALIANT, 1984). Usualmente, o desempenho de um classificador tende a diminuir a partir de uma determinada quantidade de características (BISHOP, 1995), apesar delas serem potencialmente significativas para a resolução do problema.

A redução de dimensionalidade na extração de características tem duas justificativas (HAGHIGHAT; NAMJOO, 2011). A primeira justificativa é a redução do tamanho do vetor de características, propiciando uma significativa diminuição no custo computacional associado 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.

Visão Geral da Proposta

Este trabalho realiza reconhecimento facial sujeito a oclusões parciais da face, variações de iluminação e de expressões faciais considerando uma representação compacta das imagens. O processo de reconhecimento facial é constituído por seis etapas responsáveis pela obtenção de um vetor de características representativo para uma imagem facial e pela classificação dos indivíduos. Este trabalho tem como base de pesquisa a extração de características, terceiro módulo do sistema de reconhecimento facial proposto por Li e Jain (2011).

A primeira etapa, Figura 2, consiste na aplicação do banco de filtros de Gabor curvo (HWANG et al., 2015) para a extração do vetor de características da face. A partir de uma imagem inicial realiza-se a convolução com o banco de filtros, configurado com 4 variações no grau de curvatura $c = \{0; 0,05; 0,1; 0,2\}$, gerando um conjunto de magnitudes das respostas do Gabor curvo baseadas nas suas configurações de frequência e orientações.

Em seguida, na segunda etapa, as magnitudes das respostas do Gabor curvo são segmentadas em blocos não sobrepostos. O objetivo da segmentação é diminuir a

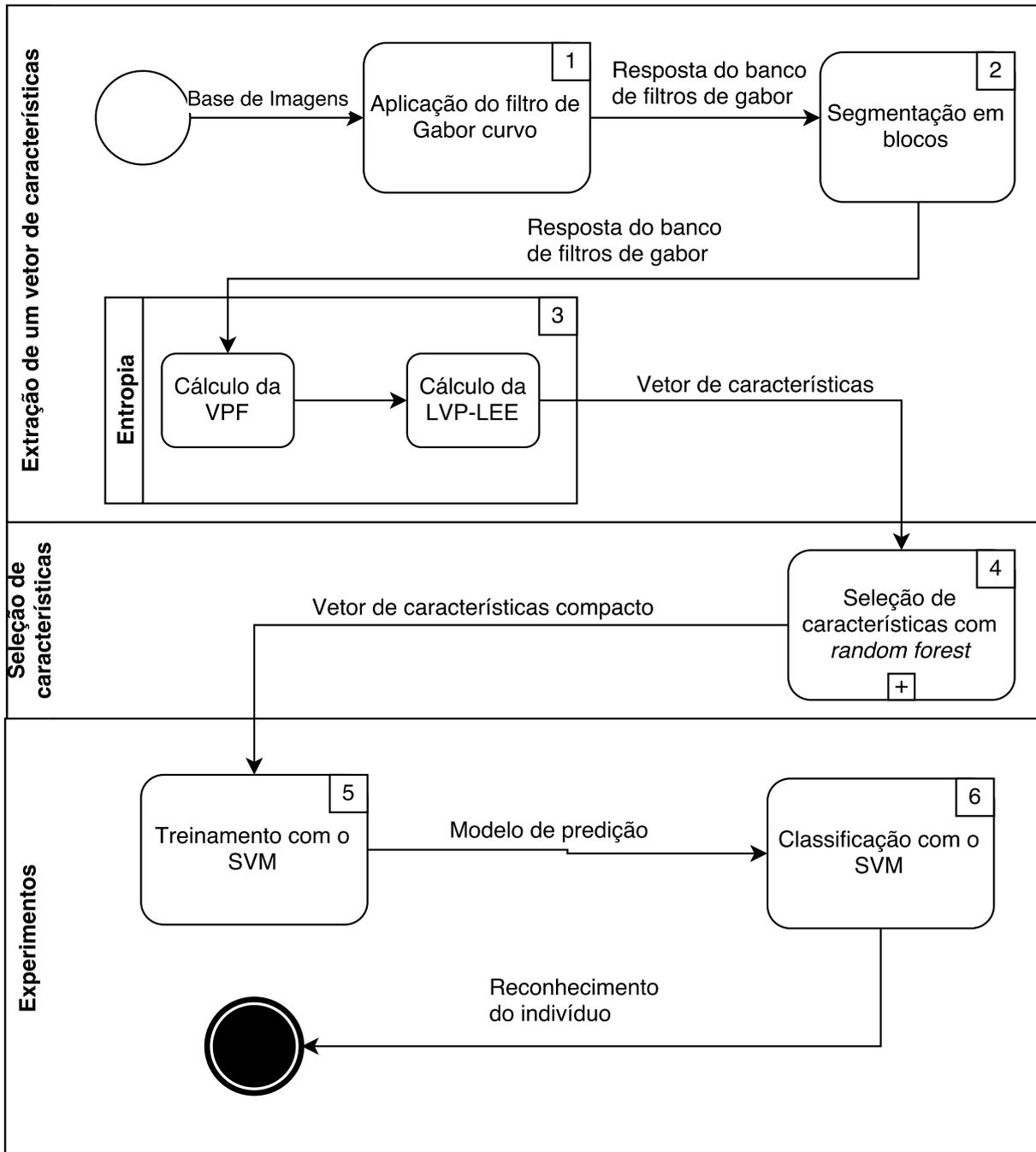


Figura 2 – Fluxograma geral da abordagem proposta.

interferência de variações locais na imagem, causadas por mudanças na iluminação ou oclusões parciais da face. Variações de iluminação podem criar sombras na face ou alterações de brilho e contraste.

Na terceira etapa, com o objetivo de manter os dados mais representativos da imagem em um vetor de dimensionalidade reduzida, utilizou-se a entropia. Devido ao fato de que a alta dimensionalidade de uma imagem original dificulta ou mesmo inviabiliza o classificador de obter bons resultados (KAR et al., 2012), tal etapa torna-se fundamental para a realização da classificação.

Como quarta etapa, realiza-se a seleção de características com o *Random Forest*. Nessa etapa os atributos mais relevantes do vetor de características são utilizados para a construção de um novo vetor compacto.

A execução do algoritmo de classificação é precedida da realização do treinamento, quinta etapa, definindo um modelo de predição que dada uma nova imagem é capaz de identificar o indivíduo.

Por fim, a representação computacional da imagem obtida na quarta etapa, é fornecida como entrada ao SVM (*Support Vector Machines*) (VAPNIK, 1995) já treinado, responsável por efetuar a classificação das imagens, reconhecendo o indivíduo.

Para a validação da abordagem proposta, comparou-se os resultados obtidos com outros de abordagens que possuem o mesmo objetivo de reconhecer faces em ambiente com oclusão, variações de iluminação e expressões faciais, são elas: SRC (WRIGHT et al., 2009), LRC (NASEEM; TOGNERI; BENNAMOUN, 2010), CESR (HE; ZHENG; HU, 2011), FSS (LEE et al., 2006), SGLasso (YUAN; LIN, 2006) e DCC (LAI et al., 2015).

Trabalhos Relacionados

Esta seção apresenta trabalhos relacionados ao problema de reconhecimento facial, realizados em faces com oclusão parcial e variações de iluminação e expressões faciais.

As abordagens SRC (*Sparse Representation-based Classification*) (WRIGHT et al., 2009) e CESR (*Correntropy-based sparse representation*) (HE; ZHENG; HU, 2011) são baseadas em representação esparsa. Nessas abordagens, ao invés de amostrar todas as características da imagem, torna-se suficiente adquirir apenas alguns componentes da informação e de maneira aleatória. Apesar de não procurar por pontos específicos como olhos, boca ou nariz, esse tipo de representação de imagem é suficiente para reconstruir a imagem.

Em (WRIGHT et al., 2009) é apresentado os principais desafios para realização do reconhecimento facial em imagens. Entre eles, as distorções ocasionadas por variações de iluminação, expressões faciais, 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 outras propostas.

Para lidar com problemas de oclusão e de ruídos no reconhecimento facial é apresentado em (HE; ZHENG; HU, 2011) um algoritmo chamado CESR (*Correntropy-based sparse representation*). Nele são combinados a representação esparsa não negativa e

o critério da máxima *correntropy*.

O problema do uso de *sparse coding* em grandes bases de imagens devido ao alto custo computacional é apresentado pela abordagem *Feature-sign search* (FSS) em (LEE et al., 2006). No trabalho 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 abordagem apresenta dificuldades sobretudo ao tratar a oclusão sobre a região dos olhos.

Os demais trabalhos que atuam como comparativo da abordagem proposta, fundamentam-se em regressão linear. LRC (*Linear Regression Classification*) (WRIGHT et al., 2009), FSS (*Feature-sign search*) (LEE et al., 2006), SGLasso (*Structured graphical lasso*) (YUAN; LIN, 2006) e DCC (*Discriminative and compact coding*) (LAI et al., 2015) consideram a imagem de teste como uma combinação linear de algumas das imagens de treinamento.

Dentre as abordagens citadas, todas apresentam maiores dificuldades para realizar o reconhecimento facial diante de oclusões contínuas nas regiões dos olhos e boca. Tal fato poderia ser atribuído a característica holística das abordagens. Na prática, o vetor de características holístico não mostra-se tão robusto a variações de iluminação, expressão facial, posição e deformações locais (WONG et al., 2012) (PIEPERS; ROBBINS, 2012) (WENG; LU; TAN, 2016). Buscando resolver essa dificuldade, nesse trabalho é proposta a utilização da entropia de blocos do 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 é segmentada em blocos não sobrepostos. Em seguida é aplicada a entropia, ressaltando características locais na representação final da imagem. Por fim é realizada a seleção de características por *random forest* de forma a diminuir a dimensionalidade da representação facial.

Objetivos

Este trabalho tem como objetivo principal a realização do reconhecimento facial sujeito a oclusões parciais da face, variações de iluminação e de expressões faciais considerando uma representação compacta das imagens.

Além disso, pretende-se alcançar os seguintes objetivos específicos:

- Realizar uma representação da face em diferentes graus de orientação, escala e translação, utilizando a Entropia de projeção do Gabor Curvo (EPGC) para extração do vetor de características da imagem;

- Reduzir a dimensionalidade do vetor de características faciais com o uso da entropia e do *random forest*;
- Verificar a eficácia da solução proposta para reconhecimento facial por meio de um conjunto de experimentos realizados nas bases de imagens *AR Face*, FERET e YALE utilizando o SVM como classificador.

Produção Bibliográfica

Os trabalhos produzidos e publicados ou aguardando publicações, relacionados a este trabalho são os seguintes:

- LIMA, EUCASSIO GONCALVES; VOGADO, L. H. S.; PASSARINHO, C. J. P.; DE ANDRADE LIRA RABELO, RICARDO. Um robusto reconhecimento facial por filtro de gabor curvo e entropia. Revista de Sistemas e Computação - RSC. , v.6, p.80 - 89, 2016.
- LIMA JUNIOR, E. G. ; RABELO, RICARDO A. L. ; PASSARINHO, C. J. P. ; VOGADO, L. H. S. . Representação compacta de face para reconhecimento sob oclusão. in: Escola Regional de Informática do Piauí (ERUPI), 2016, Teresina.
- LIMA, EUCASSIO GONCALVES; DE ANDRADE LIRA RABELO, RICARDO; PASSARINHO, C. J. P.; VOGADO, L. H. S. Reconhecimento facial sob oclusão e iluminação com uma representação compacta In: XXI Congresso Brasileiro de Automática - CBA2016, 2016, Vitória – ES. Anais do XXI Congresso Brasileiro de Automática - CBA2016. , 2016. p.1500 - 1500
- LIMA, EUCASSIO GONCALVES; DE ANDRADE LIRA RABELO, RICARDO ; PASSARINHO, C. J. P. ; VOGADO, L. H. S. . Uma abordagem baseada em entropia e filtro de gabor para um reconhecimento facial robusto. In: Simpósio Brasileiro de Automação Inteligente, 2015, Natal. ., 2015.

Contribuições

As principais contribuições relacionadas neste trabalho são elencadas a seguir:

1. A fusão de banco de filtros de Gabor curvo com a entropia da projeção de variância local e *random forest* para representação da face. Essa representação facial utilizada nos testes, apresenta significativos avanços em recentes experimentos de reconhecimento facial;

2. Determinação de uma sequência de valores que representam o grau de curvatura do banco de filtros de Gabor curvo, de modo que se obtenha os melhores valores de acurácia nos experimentos realizados;
3. Representação das imagens baseada na Entropia de projeção do Gabor Curvo para a extração de características mais representativas da face, alcançando-se resultados mais significativos que o Gabor clássico e o Gabor curvo.

Estrutura do Trabalho

O Capítulo 1 apresenta o referencial teórico do banco de filtro de Gabor, *random forest*, entropia e SVM. Esse capítulo finaliza com a exposição de trabalhos relacionados à proposição de soluções de reconhecimento facial, em ambientes com oclusões parciais da face, e variações de iluminação e expressões faciais.

O Capítulo 2 descreve a solução proposta para o problema do reconhecimento facial e apresenta cada um dos componentes utilizados no processo de representação e classificação das imagens.

O Capítulo 3 descreve os testes realizados para avaliar a solução proposta e a discussão dos resultados obtidos com os experimentos realizados.

Por fim, o Capítulo 4 apresenta as conclusões, enumera as limitações da solução proposta para representação da face e descreve as perspectivas para continuação deste trabalho.

1 Referencial Teórico

Neste capítulo são apresentados os conceitos associados aos sistemas de reconhecimento facial. Inicialmente, descreve-se o problema de extração de características de uma imagem utilizando o banco de filtro de Gabor. Em seguida, discute-se como o banco de filtros de Gabor Curvo é utilizado para extrair características mais significativas da imagem. A entropia é detalhada na sequência, como a técnica utilizada para redução de dimensionalidade do vetor de características e o *Random Forest* como ferramenta para seleção de características. Com relação à classificação das imagens, são apresentados o histórico e o funcionamento básico do SVM. Por fim, são listados trabalhos relacionados a abordagem proposta.

1.1 Representação por Gabor

A extração de características tem como objetivo representar partes importantes de uma imagem de forma eficiente, constituindo um vetor de características representativo (SHEN; BAI; FAIRHURST, 2007). Esse método é utilizado quando uma imagem é de dimensão elevada e torna-se necessária sua representação como um vetor reduzido de forma a concluir rapidamente atividades como reconhecimento.

Deteção e extração de características bem como rastreamento são usualmente combinados para resolver problemas comuns da visão computacional, dentre eles: deteção e reconhecimento de objetos, deteção e reconhecimento de faces e classificação de texturas (MESHGINI; AGHAGOLZADEH; SEYEDARABI, 2013).

Os filtros de Gabor têm sido usados amplamente em análise de padrões de aplicações. Dentre as mais importantes características da magnitude do filtro de Gabor, estão sua invariância a iluminação, rotação, escala, e translação (KAMARAINEN; KYRKI; KÄLVIÄINEN, 2006). Além disso, ele é pouco afetado por distúrbios de fotografia, como mudanças de iluminação e ruído de imagem (KAMARAINEN; KYRKI; KÄLVIÄINEN, 2006).

Proposto inicialmente em (GABOR, 1946) para processamento de sinais em uma dimensão, os filtros de Gabor foram estendidos para duas dimensões em (DAUGMAN, 1985). Dado uma imagem qualquer esta pode ser representada por um conjunto de características obtidas pelas respostas do filtro. Sendo esta a combinação de funções de Gabor de diferentes frequências.

Assim, bancos de filtros de Gabor tornam-se uma poderosa ferramenta para extração de características faciais (CHAI et al., 2014; REN et al., 2014). Eles podem extrair

características locais de uma frequência de banda limitada da imagem. O Filtro de Gabor 2D pode ser expresso segundo a Equação 1.1 (SHEN; BAI; FAIRHURST, 2007):

$$\begin{aligned} \varphi_{\Pi(q,\theta,\gamma,\eta)}(x,y) &= \frac{q^2}{\pi\gamma\eta} \exp(-(\alpha^2 x'^2 + \beta^2 y'^2)) \exp(j2\pi q x') \\ x' &= x \cos \theta + y \sin \theta \\ y' &= -x \sin \theta + y \cos \theta \end{aligned} \quad (1.1)$$

na qual q é a frequência no plano de onda, θ é o ângulo de rotação, medido no sentido anti-horário, entre a direção de propagação do plano de onda e o eixo x , α e β definem a abertura do envelope Gaussiano no maior eixo paralelo à onda e menor eixo perpendicular à onda. Os parâmetros $\gamma = \frac{q}{\alpha}$ e $\eta = \frac{q}{\beta}$ são definidos para manter a proporção constante do envelope gaussiano. Em trabalhos da literatura (WISKOTT et al., 1997; LIU; WECHSLER, 2002a), os autores comumente definem os parâmetros e utiliza-se os valores $\gamma = \eta = \sqrt{2}$. Outro parâmetro fundamental para aplicação do banco de filtros de Gabor é q_{max} que é normalmente definido como $q_{max} = 0,25$ já que informações úteis de faces estão contidas em bandas de baixa frequência (SHEN; BAI; FAIRHURST, 2007).

Para extrair características de uma imagem, um conjunto de filtros de Gabor é utilizado com diversas frequências e orientações para que se possa extrair atributos relevantes de diferentes ângulos e rotações. Cada filtro, representa uma função gaussiana modulada por uma plano de onda complexo cuja frequência e orientações são definidos pelos parâmetros q_u e θ_v , respectivamente. Segundo a literatura, é comum utilizar o banco de filtros de Gabor formado com 5 frequências e 8 orientações (CAMENT et al., 2015):

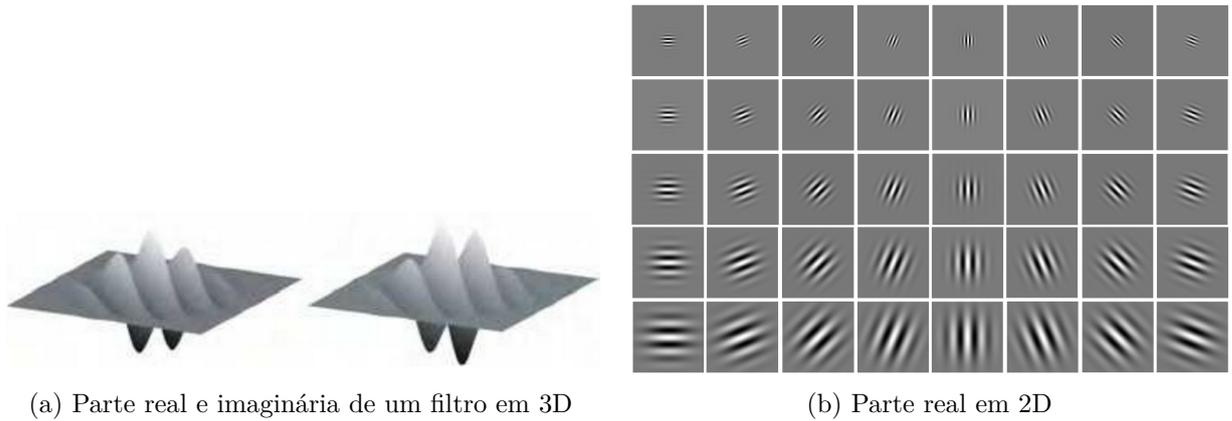
$$\begin{aligned} \varphi_{u,v} &= \varphi_{\Pi(q_u,\theta_v,\gamma,\eta)}, \\ q_u &= q_{max}/\sqrt{2^u}, \theta_v = \frac{v}{8}\pi \\ u &= 0, \dots, U-1, v = 0, \dots, V-1 \end{aligned} \quad (1.2)$$

A Equação 1.1 possui duas partes, uma real (R) e outra imaginária (I) que podem ser definidas por:

$$\varphi(x,y) = R_{(x,y)} + jI_{(x,y)} \quad (1.3)$$

Um exemplo da parte real e imaginária de um banco de filtros de Gabor pode ser visualizada em 3D na Figura 3a. A parte real do banco de filtros é apresentada em 2D na Figura 3b, na qual verifica-se o aspecto multiescalar do filtro.

Considera-se $f(x,y)$ como a escala de cinza de uma imagem de face de tamanho $a \times b$ pixels, e que $\psi_{u,v}(x,y)$ é a representação de um banco de filtros de Gabor, no qual tem-se a frequência por q_u e a orientação θ_v . O procedimento de extração de características



(a) Parte real e imaginária de um filtro em 3D

(b) Parte real em 2D

Figura 3 – Exemplos de banco de filtros de Gabor.

pode ser então descrito como a convolução da imagem da face $f(x, y)$ pelo banco de filtros de Gabor $\psi_{u,v}(x, y)$ (REN et al., 2014), conforme abaixo:

$$G_{u,v}(x, y) = f(x, y) * \psi_{u,v}(x, y). \quad (1.4)$$

Na Equação 1.4, $G_{u,v}(x, y)$ representa o resultado complexo da convolução. Por meio dele obtêm-se a magnitude das respostas ($|\varphi(x, y)|$) bem como fase (ϕ), expressas nas Equações 1.5 e 1.6 respectivamente. Na Figura 4 é apresentado um exemplo da magnitude do banco de filtro de Gabor, gerado a partir de uma imagem de uma face.

$$|\varphi(x, y)| = |R_{(x,y)}^2 + I_{(x,y)}^2|^{1/2}. \quad (1.5)$$

$$\phi(x, y) = \arctan(I_{(x,y)}/R_{(x,y)}) \quad (1.6)$$

O reconhecimento facial destaca-se como uma das mais importantes aplicações do Filtro de Gabor (LIU; WECHSLER, 2002b). Quando é realizada a convolução da representação da imagem facial com um conjunto de filtro de Gabor, a resultante é um conjunto de características mais significativas da imagem que está qualificada para ser aplicada numa técnica de reconhecimento.

1.1.1 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 têm formas curvas. Assim, infere-se que um filtro de Gabor curvo é mais apropriado do que o clássico Gabor (TORKHANI et al., 2017). Conforme é detalhado na Figura 5 na qual ao se sobrepor o filtro ao formato das linhas

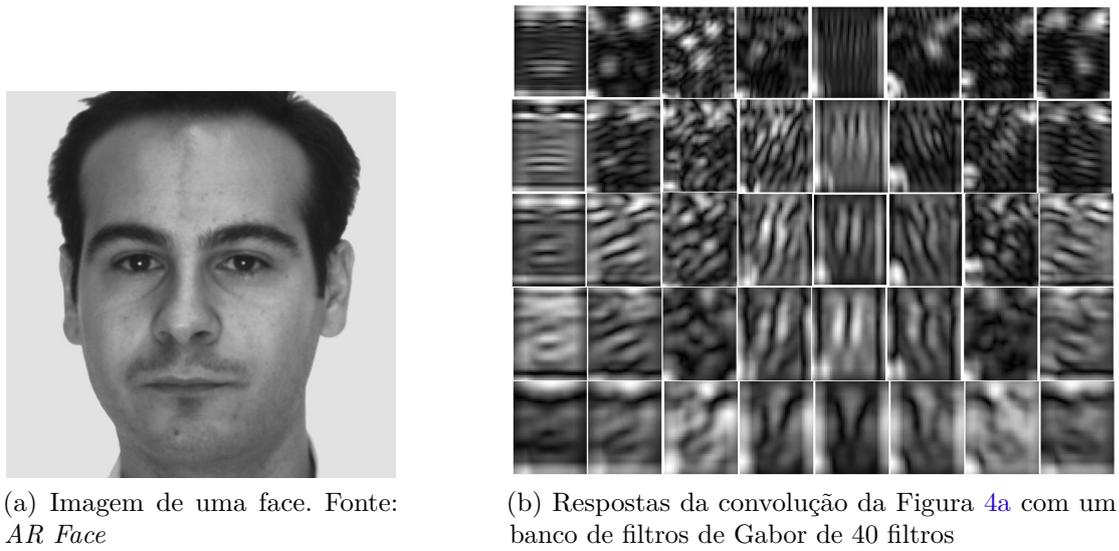


Figura 4 – Exemplo da magnitude das respostas do Gabor.

do rosto é possível deduzir que é possível extrair dados mais representativos da imagem (HWANG et al., 2015).

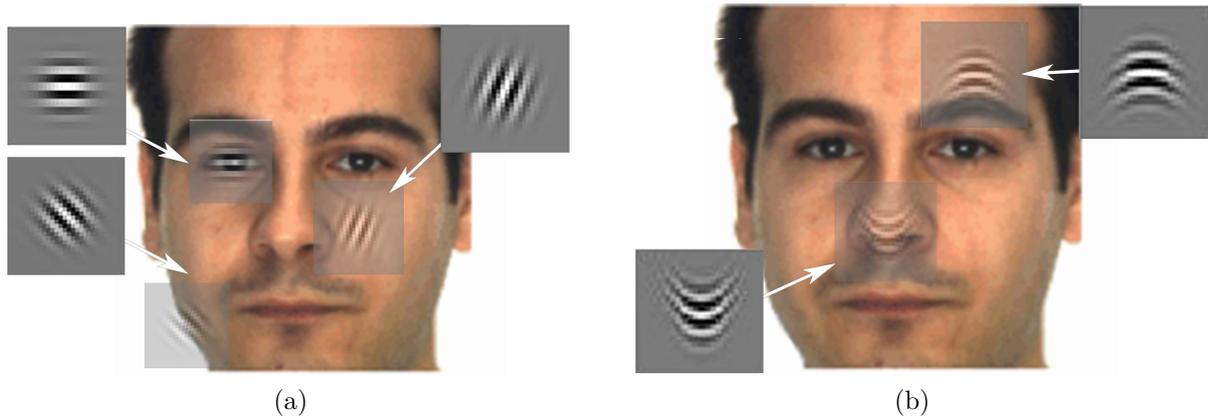


Figura 5 – Filtro de Gabor clássico (a) e o filtro de Gabor curvo (b), aplicado nos olhos nariz e bochechas.

A partir da Equação 1.1, obtem-se a Equação 1.7 representando o filtro de Gabor curvo, na qual c indica o grau de curvatura:

$$\begin{aligned} \varphi_{k,\sigma,v,\mu,c}(x,y) &= \frac{k^2}{\sigma^2} \exp\left(-\frac{k_v^2}{2\sigma^2}((x' + c(y')^2)^2 + (y')^2)\right) \\ &\times [\exp(ik_v(x' + c(y')^2)) - \exp(-\frac{\sigma^2}{2})], \end{aligned} \quad (1.7)$$

$$\begin{aligned} x' &= x \cos \varphi_\mu + y \sin \varphi_\mu \\ y' &= -x \sin \varphi_\mu + y \cos \varphi_\mu \end{aligned}$$

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 clássico, o filtro curvo utiliza 16. Tal fato, deve-se a assimetria do banco de filtros curvo, detalhado na Figura 6, que para estender-se sobre toda as orientações tendo $\varphi_\mu = \mu(\pi/8)$ necessita do dobro de filtros. A Figura 6 detalha as características multiescalares do banco de filtro utilizados.

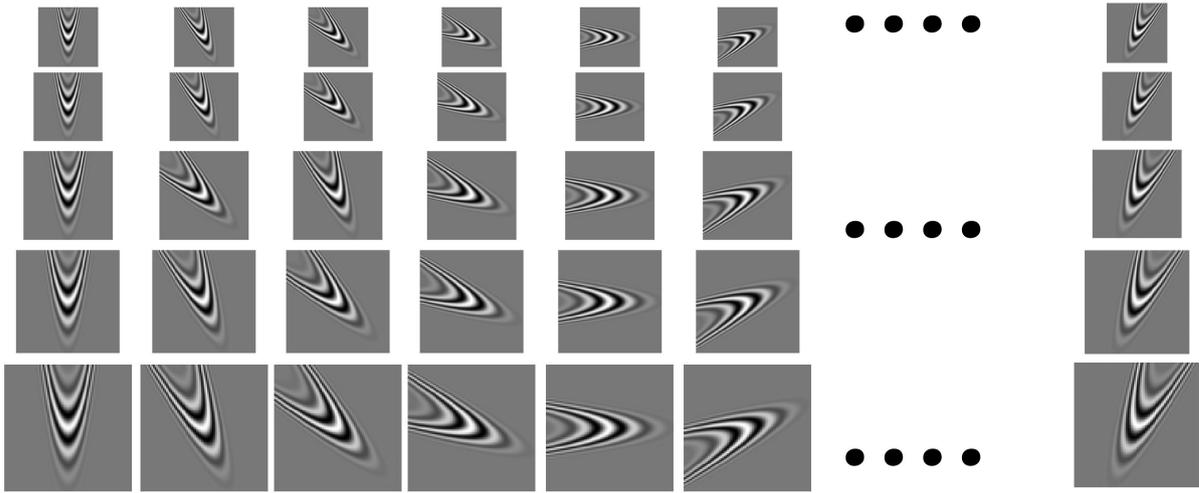


Figura 6 – Banco de filtro de Gabor curvo.

Dessa forma, é produzido um novo banco de filtros de Gabor, conforme Figura 7 que contrapõe o banco de filtro de Gabor clássico, com sua variação de curvatura $c = \{0, 05; 0, 1; 0, 2\}$. Para a definição dos valores do conjunto c , realizou-se uma sequência de experimentos, variando de $c = 0$ até 0,5. Para $c > 0, 2$, o banco de filtro de Gabor curvo apresenta uma deformidade acentuada, isso reduz a qualidade do vetor de características gerado, devido a uma menor quantidade de informações serem extraídas por meio da convolução.

1.2 *Random Forest*

Random Forest é um algoritmo de aprendizado de máquina supervisionado utilizado para classificação e regressão de dados (BREIMAN, 2001). Para efetuar a classificação, o algoritmo a partir de amostras aleatórias do conjunto de dados analisado, cria diversas árvores de decisão. Cada árvore elege individualmente uma classe específica, a qual deverá pertencer o vetor de características. A classe mais votada por todas as árvores é a classificação obtida pelo *Random Forest*.

Uma das principais características do *Random Forest* é a sua capacidade de medir o grau de importância de cada característica para a predição da classe.

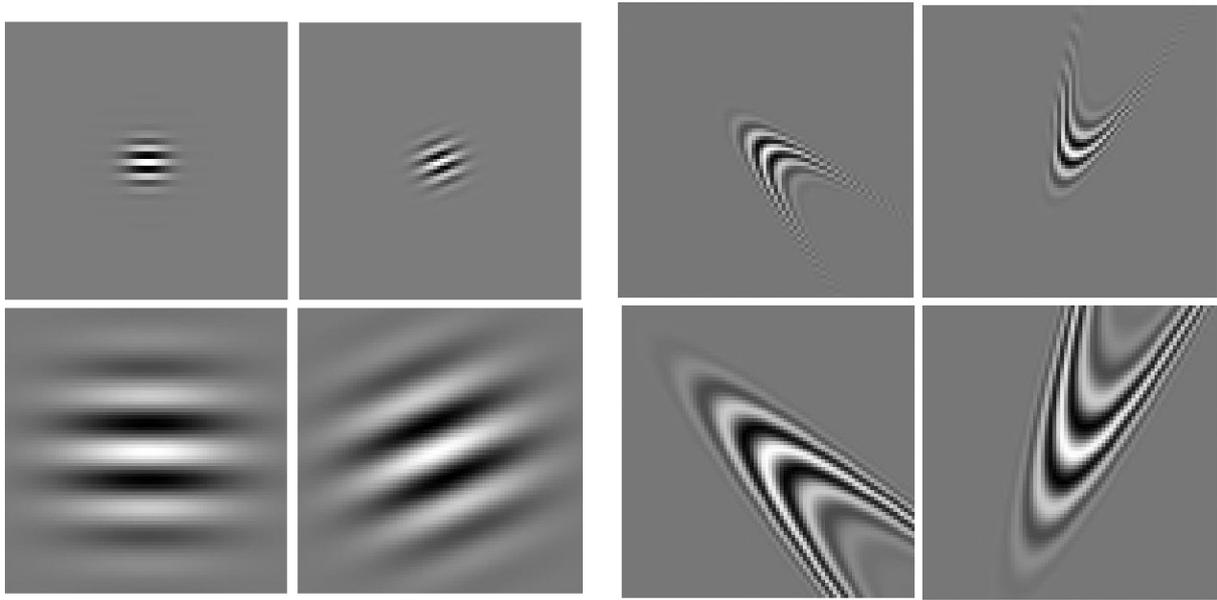


Figura 7 – Banco de filtro de Gabor e banco curvo.

O número de árvores na floresta é uma das variáveis principais a serem ajustadas quando da execução do *Random Forest*. Aumentando-se o seu valor, melhora-se o resultado obtido assim como o custo computacional. Contudo, a melhoria diminui à medida que o número de árvores aumenta. Portanto, existirá um certo ponto que o benefício no desempenho da predição com um maior número de árvores será menor que o tempo computacional exigido graças ao acréscimo.

1.3 *Local Variance Projection Log Energy Entropy Features*

1.3.1 Função de projeção de variância

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 tolerantes a mudanças de iluminação (ZHOU; GENG, 2004) (FENG; YUEN, 1998), 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 variância vertical e horizontal da imagem é definida como:

$$VPF_v(x) = \frac{1}{y_2 - y_1} \sum_{y=y_1}^{y_2} [f(x, y) - IPF_v(x)]^2, \quad (1.8)$$

$$VPF_h(y) = \frac{1}{x_2 - x_1} \sum_{x=x_1}^{x_2} [f(x, y) - IPF_h(y)]^2, \quad (1.9)$$

onde $VPF_v(x)$ e $VPF_h(y)$ são as funções 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:

$$IPF_v(x)' = \frac{1}{y_2 - y_1} \sum_{y=y_1}^{y_2} f(x, y), \quad (1.10)$$

$$IPF_h(y)' = \frac{1}{x_2 - x_1} \sum_{x=x_1}^{x_2} f(x, y), \quad (1.11)$$

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

1.3.2 Entropia

O conceito de entropia, introduzido por Shannon em (SHANNON, 1948) 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 (CHANG et al., 2006). 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:

$$S = - \sum_{i=1}^k p_i \cdot \log p_i \quad (1.12)$$

Uma imagem é uma fonte de informação, tal qual um sinal unidimensional, portanto a imagem pode ser representada por meio de sua entropia. Para uma imagem $f(i, j)$ de tamanho $M \times N$, o valor da entropia de Shannon de uma imagem é definido como:

$$En = - \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p(i, j) \log p(i, j) \quad (1.13)$$

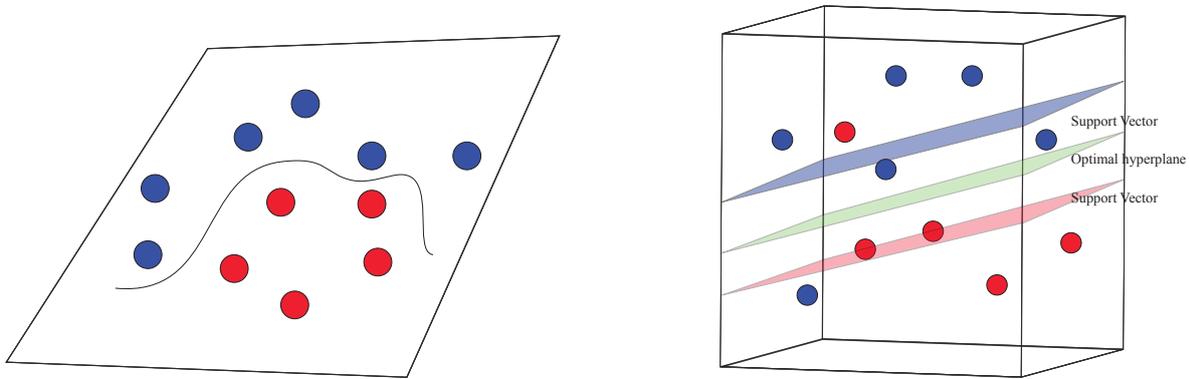
onde $p(i, j) = f(i, j) / \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i, j)$, e $f(i, j)$ é a intensidade da escala de cinza.

1.4 SVM (Máquinas de Vetores de Suporte)

Máquina de Vetores de Suporte (*Support Vector Machine*) são uma poderosa técnica para classificação de padrões (VAPNIK, 1995). O SVM cria um hiperplano ou um conjunto de hiperplanos em um alto espaço dimensional, que pode ser usado para classificação por meio do aprendizado supervisionado.

O SVM efetua o reconhecimento de padrões entre duas classes construindo esse plano ou hiperplano que maximiza a distância entre dois pontos próximos no conjunto dos pontos de treinamento (HEISELE; HO; POGGIO, 2001), conforme Figura 8. Informalmente, um hiperplano pode ser definido como um subespaço com uma dimensão a menos que o espaço de configuração do ambiente.

De um conjunto de pontos $x_i \in R^2$, $i = 1, 2, \dots, N$ 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 8 – Hiperplano gerado pelo SVM, maximizando a distância entre os dados de treinamento.

Esse hiperplano ótimo é definido por:

$$h(x) = \sum_{i=1}^l \alpha_i y_i x_i \cdot x + b, \quad (1.14)$$

Um novo ponto x é classificado computando a Equação 1.15 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.

$$d(x) = \frac{\sum_{i=1}^l \alpha_i y_i x_i \cdot x + b}{\|\sum_{i=1}^l \alpha_i y_i x_i\|}, \quad (1.15)$$

As definições representadas pelas Equações (1.14) e (1.15) 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 $z = \Phi(x)$ pertencente a um espaço dimensional superior no qual os dados são separáveis por um hiperplano. Uma propriedade

fundamental nessa estrutura é que o produto $\Phi(x) \cdot \Phi(y)$ pode ser reescrito como uma função *kernel* $K(x, y)$. A superfície de decisão é estabelecida pela Equação (1.16):

$$m(x) = \sum_{i=1}^l \alpha_i y_i K(x, x_i) + b, \quad (1.16)$$

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 um única classe de todas as outras classes restantes (SCHÖLKOPF; BURGESS; VAPNIK, 1995) (CORTES; VAPNIK, 1995).

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 trabalho, a biblioteca libSVM (CHANG; LIN, 2011), adota o mecanismo proposto inicialmente para treinamento em par, de redes neurais detalhado em Knerr, Personnaz e Dreyfus (1990) e aplicada posteriormente para SVMs por Kressel (1999).

1.5 Considerações Finais

Neste capítulo, foram descritas as técnicas matemáticas utilizadas na implementação da abordagem proposta. Destaca-se o banco de filtros de Gabor curvo para a extração de características em conjunto com a entropia e o *Random Forest*. Em seguida, apresentou-se o SVM como classificador a ser utilizado, com suas particularidades e configurações. Por fim, foram relatados algumas abordagens propostas na literatura que atuam no problema de reconhecimento facial.

2 Abordagem Proposta

Neste capítulo é apresentada a abordagem proposta nesse trabalho para um reconhecimento facial. Explicita-se, cada uma das metodologias aplicadas para a representação da imagem da face (magnitude do banco de filtros de Gabor curvo, segmentação das respostas em blocos não sobrepostos, cálculo variância, entropia e *Random Forest*), bem como os parâmetros dos experimentos realizados.

2.1 Componentes da Abordagem

A abordagem proposta visa resolver o problema do reconhecimento facial considerando oclusões parciais da face, variações de iluminação e de expressões faciais. Para isso, a abordagem aplica o banco de filtro de Gabor curvo, em conjunto com a técnica de projeção local de variância *log energy entropy*, de modo a extrair o vetor de características, e posteriormente realizar a redução do vetor. Dentre os três grandes grupos nos quais os algoritmos de reconhecimento facial são classificados (tradicionais, heterogêneos e homogêneos) (HWANG et al., 2015), a abordagem proposta nesse trabalho, pertence ao grupo dos homogêneos, extraindo o máximo de um algoritmo, variando os seus parâmetros.

A proposta apresentada neste trabalho é dividida em 6 etapas principais: aplicação do filtro de Gabor curvo, segmentação em blocos, uso da entropia, seleção de características com *random forest*, treinamento e classificação com o SVM. As etapas e suas interações são descritas na Figura 9 e detalhadas a seguir.

2.1.1 Aplicação do banco de filtros de Gabor curvo

O filtro de Gabor tem sido usado amplamente em análise de padrões de aplicações (LIU; WECHSLER, 2002b). 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 computacional inicial da imagem.

Na primeira etapa, gerou-se um banco de filtros de Gabor curvo, com o número de frequências (u) e orientações (v) de acordo com a Equação 2.1, na qual o número de orientações varia conforme o grau de curvatura de c , em concordância com trabalhos estado da arte (SHEN; BAI; FAIRHURST, 2007; HAGHIGHAT; NAMJOO, 2011).

Para o grau de curvatura foi definido o conjunto $c = \{0; 0,05; 0,1; 0,2\}$ que produz um banco de filtro de Gabor curvo. Para determinar os valores de c , efetuou-se um

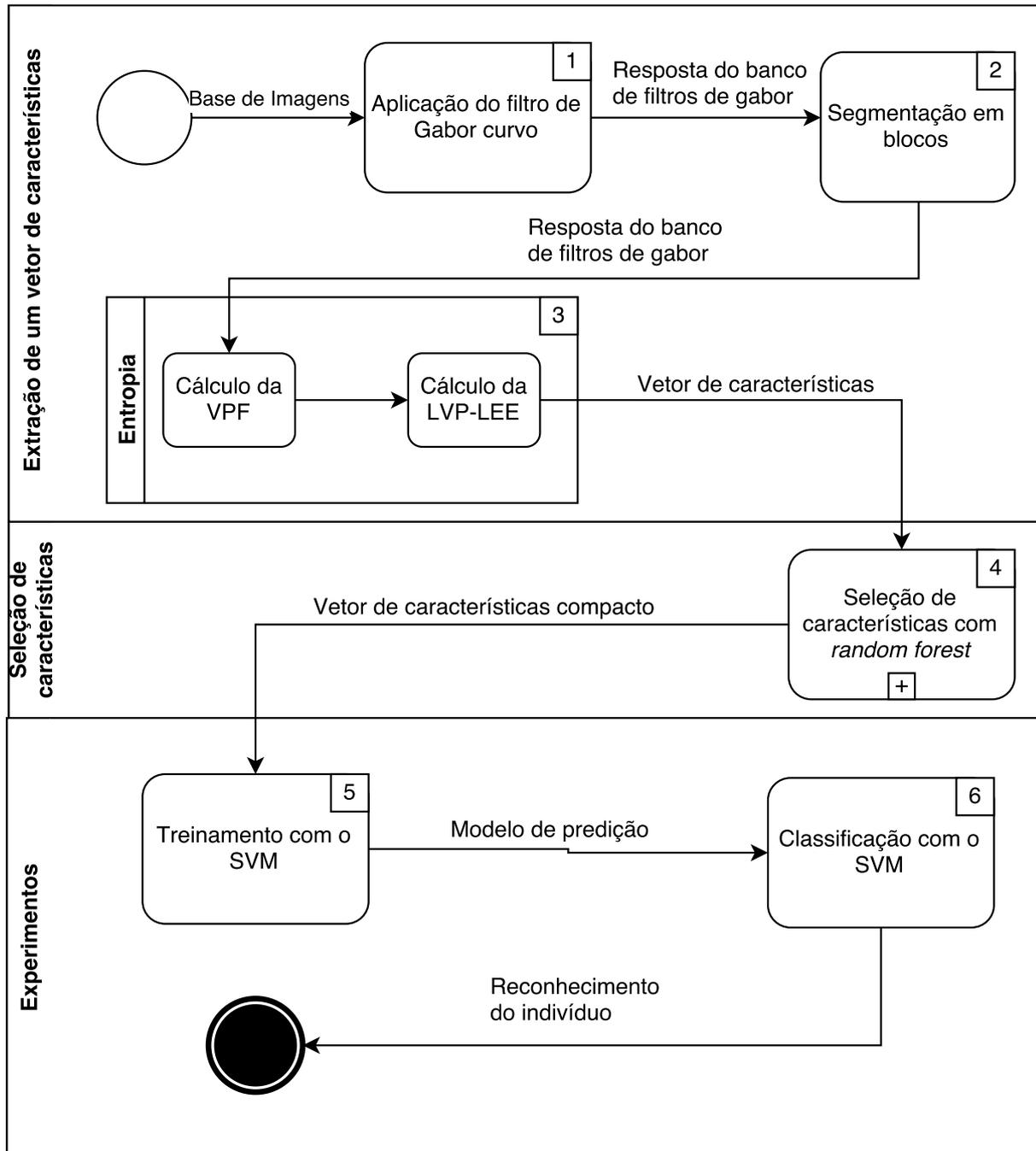


Figura 9 – Fluxograma geral da abordagem proposta.

conjunto de experimentos, variando de $c = 0$ até 0,5. Para $c > 0,2$, o banco de filtro de Gabor curvo apresenta uma deformidade acentuada, reduzindo a qualidade do vetor de características gerado. Caso c seja diferente de zero (0) o número de orientações é 16, o dobro da quantidade do filtro não curvo, com 8 orientações, devido às características

assimétricas do filtro. Totalizam-se assim, $5 \times 8 + 8 \times 16 \times 3 = 424$ filtros,

$$V = \begin{cases} 8 & \text{se } c = 0 \\ 16 & \text{se } c \neq 0 \end{cases} \quad (2.1)$$

Com o banco de filtros de Gabor curvo gerado, é efetuada a convolução de cada um dos filtros pertencente ao banco pela imagem de entrada. Assim, obtêm-se a geração da magnitude das respostas dos bancos de filtro de Gabor curvo, constituindo o conjunto de características iniciais da face. Essa coleção das respostas do Gabor curvo é o produto final da Etapa 1, conforme pode ser visto na Figura 9. Como forma de otimizar a aplicação do filtro, obtendo-se tempos menores de execução, a convolução foi efetuada por intermédio de GPU (PANG; CHOI; QIN, 2016). Utilizou-se a implementação do Matlab (THE MATHWORKS, INC., 2014) para manipulação de placas gráficas, abstraindo a complexidade de desenvolvimento da implementação nativa das unidades.

2.1.2 Segmentação em blocos

Na etapa 2 é efetuado um processamento com o objetivo de preparar os dados para a utilização da entropia e conferir maior robustez a oclusões parciais da face. 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 (ZHANG; CHEN, 2008). Uma abordagem alternativa para resolução desse problema é a segmentação da magnitude das respostas do banco de filtro de Gabor curvo, obtidas ao fim da etapa 1, em pequenos blocos não sobrepostos. A partir dessa segmentação, cada bloco de uma mesma imagem original possuirá distribuição de intensidades diferentes. Computacionalmente, essa técnica de segmentação é rápida, simples e não se baseia na detecção de importantes áreas da face como olhos, nariz e boca (CHAN et al., 2015).

Considerando uma imagem $I(x, y)$ com dimensões $M \times N$, a magnitude das respostas do Gabor curvo da face é primeiramente segmentada em L pequenos blocos não sobrepostos. O tamanho de cada bloco representado por $f''(x, y)$ é dado por $M/\sqrt{L} \times N/\sqrt{L}$, e é dependente do tamanho da imagem da face utilizada. Para o caso das bases de imagens utilizadas nesse trabalho, houve a padronização do tamanho das imagens obtendo-se $L = 256$, $M = 48$ e $N = 56$. A partir de então, para a extração do vetor de características são utilizados apenas os blocos individualmente ao invés da imagem como um todo, permitindo uma maior robustez a mudanças efetuadas em outras áreas da imagem.

2.1.3 Entropia

Em seguida, tem-se a terceira etapa que consiste em utilizar a combinação da *local variance projection log energy entropy features*. Apresentada em (ZHANG; CHEN, 2008), sua finalidade é extrair características e reduzir a dimensionalidade do vetor de características obtido a partir da etapa 2.

Após a segmentação em blocos não sobrepostos da magnitude das respostas do Gabor curvo, pode-se em sequência aplicar no vetor de características alguma técnica para redução de dimensionalidade. Aqui, optou-se pela *Local Variance Projection Log Energy Entropy* (LVP-LEE) (ZHOU; GENG, 2004) (FENG; YUEN, 1998), ao mesmo tempo que extrai características, cria um vetor de características mais representativo. Para cada magnitude das respostas do Gabor curvo, aplica-se o método baseado na entropia. Ao ser aplicada a LVP-LEE em cada resposta do banco de filtros de Gabor curvo, duas novas matrizes de dados são produzidas, representando a entropia da projeção da variância sobre a direção horizontal e vertical.

Nessa etapa, há a geração do vetor de características final representativo, abstração da face original que é dada como entrada da abordagem. Tomando o intervalo $[k_1, k_2]$ como as posições dos blocos sobre a função f'' , pode-se aplicar a técnica da projeção da variância em cada um dos blocos pertencentes às linhas (U) e colunas (V) definidas da seguinte forma:

$$U[i] = \frac{1}{k_2 - k_1} \sum_{k=k_1}^{k_2} [f''(i, k) - g_v(i)']^2, 1 \leq i \leq M/\sqrt{L} \quad (2.2)$$

$$V[j] = \frac{1}{k_2 - k_1} \sum_{k=k_1}^{k_2} [f''(k, j) - g_h(j)']^2, 1 \leq j \leq N/\sqrt{L} \quad (2.3)$$

onde:

$$g_v(i)' = \frac{1}{y_2 - y_1} \sum_{y=y_1}^{y_2} f''(i, j), \quad (2.4)$$

$$g_h(j)' = \frac{1}{x_2 - x_1} \sum_{x=x_1}^{x_2} f''(i, j), \quad (2.5)$$

Em seguida, é calculada a *log energy entropy* para cada uma das linhas e colunas resultantes das Equações 2.2 e 2.3:

$$H_{RP} = - \sum_{i=1}^{M/\sqrt{L}} e_i \log(e_i) \quad (2.6)$$

$$H_{CP} = - \sum_{i=1}^{N/\sqrt{L}} e_j \log(e_j) \quad (2.7)$$

onde

$$e_i = U[i] / \sum_{i=1}^{M/\sqrt{L}} U[i] \quad (2.8)$$

$$e_j = V[j] / \sum_{j=1}^{N/\sqrt{L}} V[j] \quad (2.9)$$

Para uma única imagem qualquer, são produzidas ao final $2L$ matrizes com os valores da entropia blocada, exibidas na forma de duas imagens conforme a Figura 10. Essas características são nomeadas como *vertical variance projection entropy vector* e *horizontal variance projection entropy vector* e após serem concatenadas constituem um só vetor de características.

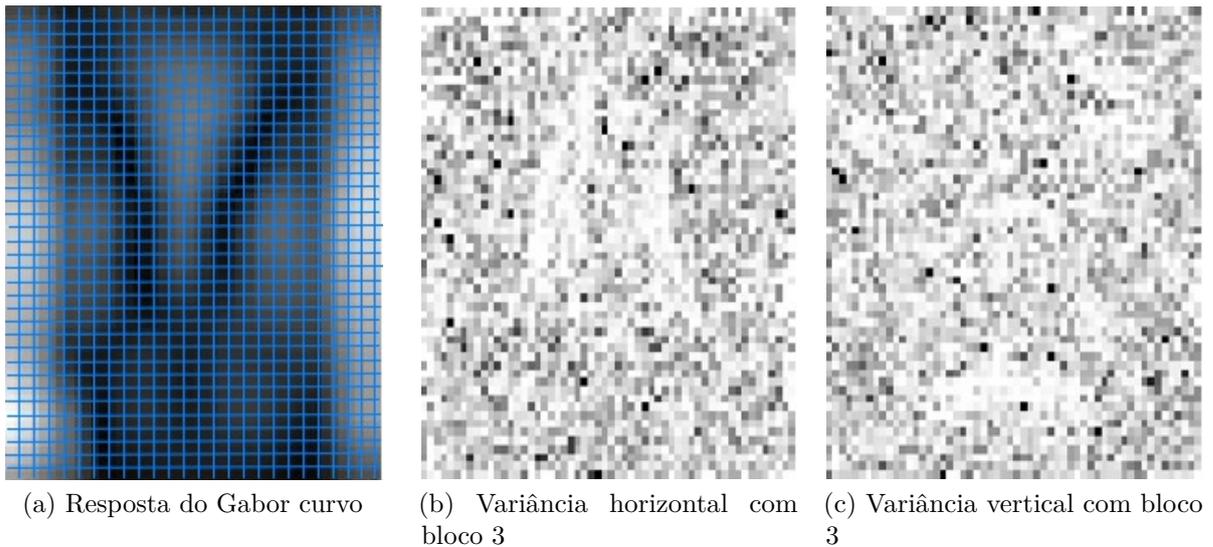


Figura 10 – Cada bloco da resposta da magnitude do Gabor curvo tem a abordagem baseada em entropia calculada, resultando em duas novas representações (Etapa 3).

2.1.4 Seleção de características por *Random Forest*

Um grande número de características é fornecido por meio da extração de características faciais. Dessa forma é necessário reduzir o tamanho do vetor de características utilizado da representação da imagem facial. Para isso é necessário identificar os elementos mais significativos do vetor, fazendo uso de uma técnica de seleção de características.

Dentre vários algoritmos que fornecem a relevância de cada elemento para a classificação foi utilizado neste trabalho o *Random Forest*. Sua escolha deve-se ao seu

amplo uso na seleção de variáveis em problemas que envolvem grande quantidade de dados (JANITZA; TUTZ; BOULESTEIX, 2016).

Na Etapa 4 é realizada a seleção de características de forma paralela, conforme representado na Figura 11. O vetor de características V , provido pela Etapa 3 é dividido em s subvetores de tamanho t . Os k elementos de cada subvetor mais relevantes para a classificação são selecionados por meio do *Random Forest*. Nesse processo definiu-se o número de árvores para 250, conforme testes efetuados com o parâmetro variando em 1, 25, 50, 250 e 1000. Na Figura 12, verifica-se que não há um crescimento relevante com valores acima de 250 para o número de árvores. Os elementos então selecionados em cada subvetor são concatenados, formando um novo vetor candidato que é submetido ao mesmo procedimento descrito. Assim, na segunda iteração é obtido um novo vetor, representação final do indivíduo.

As escolhas dos valores dos parâmetros s e k são de fundamental importância para a obtenção de um vetor de características relevante. Para a melhor escolha desses valores, foi efetuado um conjunto de experimentos em bases faciais. Os experimentos tinham por princípio modificar os valores de s e k e em seguida executar a Etapa 4 da abordagem (Figura 11) obtendo o valor da acurácia. A execução do procedimento tem como resultado a Tabela 1. Por meio dela, considerando a maximização da acurácia e a minimização do tamanho do vetor de características, identifica-se que para a proposta, os melhores valores são $k = 100$ e $s = 28$, totalizando um vetor de características de tamanho 2800.

As quatro etapas da abordagem proposta, já descritas, correspondem ao módulo de extração de características de um sistema de reconhecimento de face, um dos aspectos principais a serem abordados para o reconhecimento facial.

2.1.5 Treinamento e classificação com o SVM

A partir do vetor de características representativo gerado na Etapa 4, inicia-se o processo de treinamento do SVM. Na etapa 5, a matriz de características representando os diversos indivíduos da base de imagens é submetida ao classificador utilizado, o SVM. Após o treinamento, o SVM, gera como resultado, um modelo de predição de dados para classificação das imagens.

Por fim, na sexta etapa, um vetor gerado pelas etapas de 1 a 4 é fornecido ao SVM para que ele efetue a classificação do vetor representativo de uma face e consequente identificação do indivíduo. Como ferramenta de apoio a execução do SVM como classificador, utilizou-se a biblioteca LIBSVM (CHANG; LIN, 2011), com o *kernel Radial basis function*. Nessa etapa, utilizou-se a técnica *cross-validation*, utilizada para estimar erro de generalização em classificadores. Nessa técnica, divide-se o conjunto de dados de treinamento em r subconjuntos com valores não sobrepostos e de mesmo tamanho. As

SVM são submetidas aos $r - 1$ subconjuntos gerados, de forma que seja obtido o modelo de classificação e para a validação usa-se o último subconjunto. Esse processo é executado r vezes e assim cada subconjunto é empregado para validação uma vez. Ao final da execução da técnica é obtida a acurácia da abordagem para reconhecimento facial bem como os parâmetros utilizados no SVM.

Dentre os parâmetros principais do SVM para o *kernel* RBF utilizado, encontram-se C e γ . A determinação dos melhores valores para esses parâmetros é fundamental para a melhoria da acurácia no processo de classificação (LIAO et al., 2015). O parâmetro C controla a compensação dos erros cometidos na fase de treinamento, permitindo que alguns erros de classificação sejam cometidos, enquanto o γ controla o raio da função do *kernel*. Na abordagem proposta, ao final do *cross-validation* encontra-se os melhores valores para os seguintes parâmetros do SVM: a penalidade $C = 8$ e o parâmetro $\gamma = 3,0517578125 \cdot 10^{-5}$ do kernel RBF.

2.2 Considerações Finais

Neste capítulo foram apresentados detalhes da abordagem proposta para o reconhecimento facial e seus componentes. O capítulo descreveu a interação entre os componentes de forma a elucidar como eles colaboram para definir o vetor de características final, que representa a face de um indivíduo.

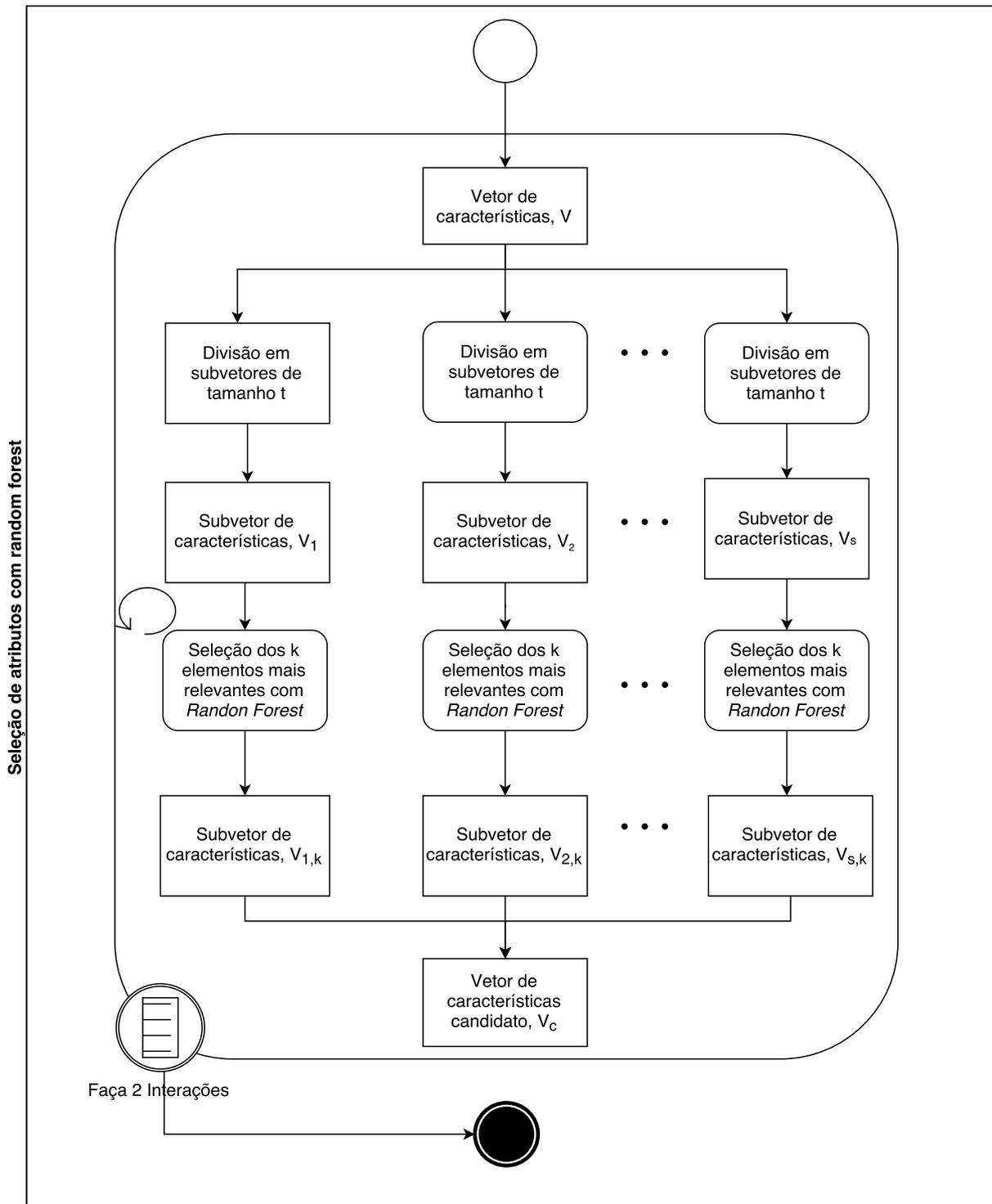


Figura 11 – Seleção de características com *random forest* (Etapa 4).

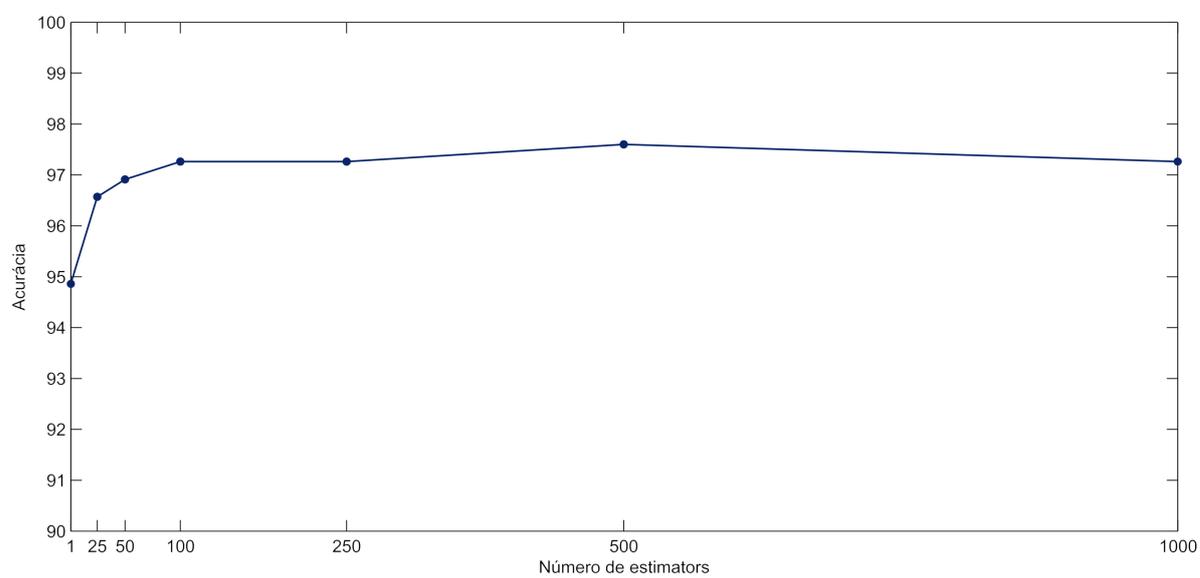


Figura 12 – Variação da acurácia pelo número de árvores para a FERET.

Acurácia	k	s
0.713675213675	100	32
0.705128205128	100	40
0.700854700855	150	24
0.696581196581	100	35
0.685897435897	250	15
0.685897435897	200	20
0.683760683761	250	16
0.683760683761	100	30
0.681623931624	100	28
0.67094017094	350	10
0.668803418803	250	14
0.666666666667	200	16
0.666666666667	150	20
0.662393162393	650	6
0.662393162393	200	15
0.660256410256	400	10
0.660256410256	250	12
0.660256410256	200	14
0.65811965812	150	21
0.651709401709	600	6
0.649572649573	300	10
0.647435897436	750	5
0.645299145299	550	6
0.645299145299	300	12
0.643162393162	700	5
0.641025641026	650	5
0.636752136752	500	6
0.634615384615	800	5
0.630341880342	550	7
0.630341880342	250	10
0.626068376068	1300	3
0.626068376068	400	8
0.623931623932	1200	3
0.623931623932	900	4
0.623931623932	600	5
0.621794871795	1150	3
0.617521367521	150	15
...
0.356837606838	100	5
0.331196581197	150	3
0.32264957265	100	4
0.305555555556	250	2
0.269230769231	200	2

Tabela 1 – Busca pelos melhores valores de s e k .

3 Resultados e Discussões

Nesta capítulo são apresentados os detalhes da aplicação da abordagem nas bases *AR Face Database* (MARTINEZ, 1998), FERET (PHILLIPS et al., 2000), YALE (GEORGHIADES; BELHUMEUR; KRIEGMAN, 2001) e os resultados obtidos. 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 (WRIGHT et al., 2009), LRC (NASEEM; TOGNERI; BENNAMOUN, 2010), CESR (HE; ZHENG; HU, 2011), FSS (LEE et al., 2006), SGLasso (YUAN; LIN, 2006) e DCC (LAI et al., 2015) que foram utilizados para comparar a abordagem proposta neste trabalho.

Os experimentos foram realizados a partir de um protótipo implementado em Matlab e *Scikit-learn* (PEDREGOSA et al., 2011). A versão do Matlab utilizada foi a R2014b de 64 bits enquanto que o *Scikit-learn* utiliza a 0.18, ambos executam sob o *Windows* 10, num computador Intel®Core™ i7-4510U CPU @ 2.00GHz com 8GB de memória RAM e uma placa de vídeo NVIDIA GeForce GT 740M 2GB 384 Núcleos CUDA 1.8 GHz.

Vale destacar que a convolução do banco de filtros de Gabor curvo com a imagem da face de um indivíduo, na etapa 1 da abordagem proposta, é efetuada por meio da GPU. Ainda que alcançar um reduzido tempo de execução não seja um dos objetivos da proposta apresentada nesse trabalho, consegue-se a diminuição no tempo de execução dessa etapa, reduzindo o tempo final de execução em toda a abordagem. O *Scikit-learn* é o responsável pela implementação do *Random Forest*.

3.1 AR Face

Constituída de mais de 4000 imagens coloridas de um total de 126 pessoas, a *AR Face Database* é uma importante base para verificação de acurácia em algoritmos de reconhecimento facial. 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 experimentos, as imagens foram redimensionadas para 48x56 *pixels*, de forma a aproximar-se das dimensões utilizadas pelos demais algoritmos que são usados como comparativo. A Figura 13 exibe imagens de faces com os principais tipos de ruídos da base, exemplificando

cada subconjunto que compõe a *AR Face*: Figura 13a, variação de iluminação, Figura 13b oclusão por óculos de sol, Figura 13c variação de iluminação e oclusão por óculos de sol, Figura 13d oclusão por cachecol, Figura 13e variação de iluminação e oclusão por cachecol, Figura 13f face com expressão neutra, Figura 13g raiva e Figura 13h sorriso.

A *AR Face* foi particionada 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 em diversos cenários. Além disso, por meio desse particionamento é possível realizar comparações diretas com outras representações faciais utilizadas para reconhecimento de face. A seguir cada subconjunto é detalhado:

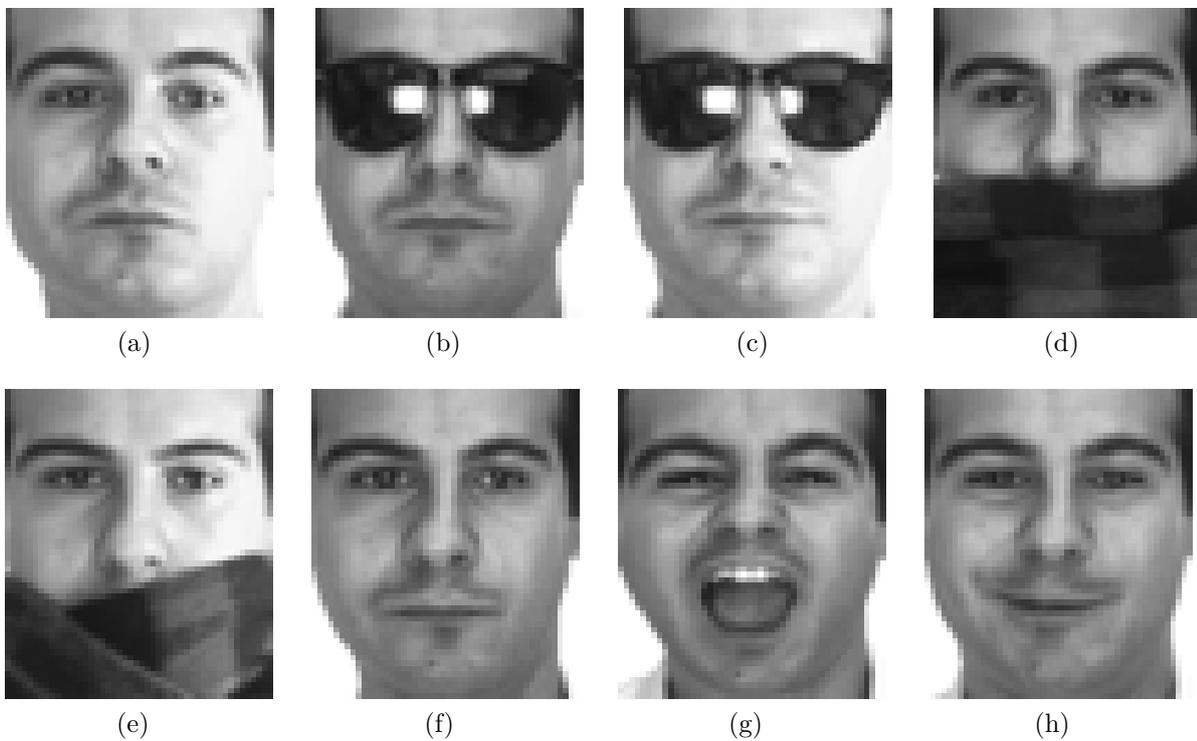


Figura 13 – Faces extraídas da *Ar Face Database*.

3.1.1 Variação apenas de iluminação

A imagem de uma face sofre efeitos bastantes 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 13a) na *AR Face* referentes a cada indivíduo compreendem apenas essa variação de iluminação. A Tabela 2 contém os valores referentes a acurácia do reconhecimento facial dos indivíduos da *AR Face*, aplicando diferentes tipos de representação e classificação das imagens. Os melhores resultados nos termos de acurácia são os obtidos por CC e DCC ($\simeq 80\%$ e $\simeq 71\%$

respectivamente). Eles representam um significativo avanço diante das demais abordagens, que alcançam o máximo de acerto de 48,74%. A abordagem proposta sobrepõe os outros métodos por percentual significativo de até 12,5%, obtendo 98% de acerto com o uso do Gabor com a entropia, Figura 14. Com o uso de Gabor curvo e Entropia atinge-se 99,83% de acerto, percentual que é melhorado com o Gabor Curvo com a entropia e *random forest* alcançando 99,50%.

Verifica-se assim que a invariância do filtro de Gabor a mudanças na iluminação é de fundamental importância para o resultado obtido. A Tabela 3 exibe os valores da precisão e *recall* para cada subconjunto. A precisão é a habilidade do classificador não rotular uma amostra negativa como positiva, enquanto que o *recall* é a capacidade do classificador encontrar todas as amostras positivas.

3.1.2 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 13b) constituem uma oclusão frontal causada por um óculos de sol, possuindo cerca de 20% da área total obstruída. De acordo com a Tabela 2, observa-se que DC e DCC apresentam os melhores percentuais, com 74,79% e 72,69% de acerto. A proposta de uso do Gabor curvo e entropia mostrou-se eficiente nesse teste obtendo 96,00% de acerto, Figura 15. 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.

3.1.3 Iluminação + óculos de sol

Esse experimento combina o desafio gerado pela variação da iluminação presente no primeiro teste, com a oclusão por óculos de sol do segundo. A Figura 13c contém exemplos extraídos das imagens 9, 10, 22 e 23 que exibem 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 representação facial com Gabor e Entropia gera um acerto de 94,75% seguido por 96,50% do Gabor curvo e entropia, evidenciando a evolução com relação às técnicas anteriormente propostas, quando há uma combinação de intensa variação de luz e oclusão, Figura 16.

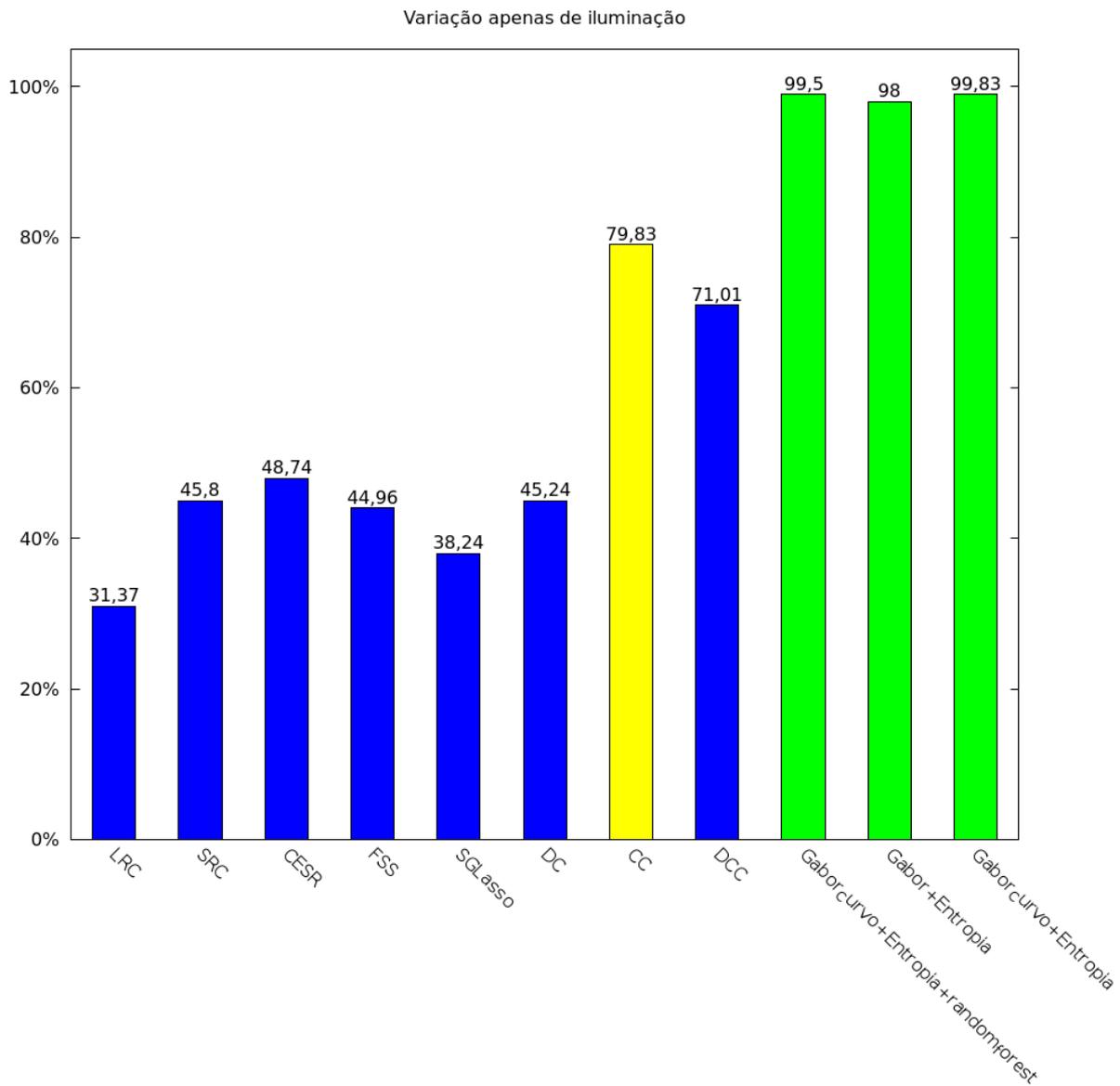


Figura 14 – Comparação de diversas abordagens com o subconjunto da *AR Face* variação de iluminação.

Abordagem	LRC	SRC	GESR	FSS	SGLasso	DC	CC	DCC	Gabor Curvo + Entropia + <i>random forest</i>	Gabor + Entropia	Gabor Curvo + Entropia
Iluminação	31,37	45,80	48,74	44,96	38,24	45,24	79,83	71,01	99,50	98,00	99,83
Oclusão por óculos de sol	25,21	28,99	68,49	28,99	21,85	74,79	3,78	72,69	96,00	76,00	78,00
Iluminação + óculos de sol	8,19	15,13	20,80	14,50	11,55	23,95	3,15	22,48	96,50	94,75	82,00
Oclusão por cachecol	94,96	95,38	96,64	95,38	93,28	97,06	68,49	97,06	96,50	84,50	86,00
Iluminação + Oclusão por cachecol	18,28	29,41	36,76	27,31	21,22	29,20	63,87	45,80	99,25	98,25	97,25
Mudança de expressão facial	x	x	x	x	x	x	x	x	99,00	98,25	89,12
Base completa	x	x	x	x	x	x	x	x	98,03	99,34	98,15

Tabela 2 – Comparação dos resultados da EPGC, com abordagens da literatura.

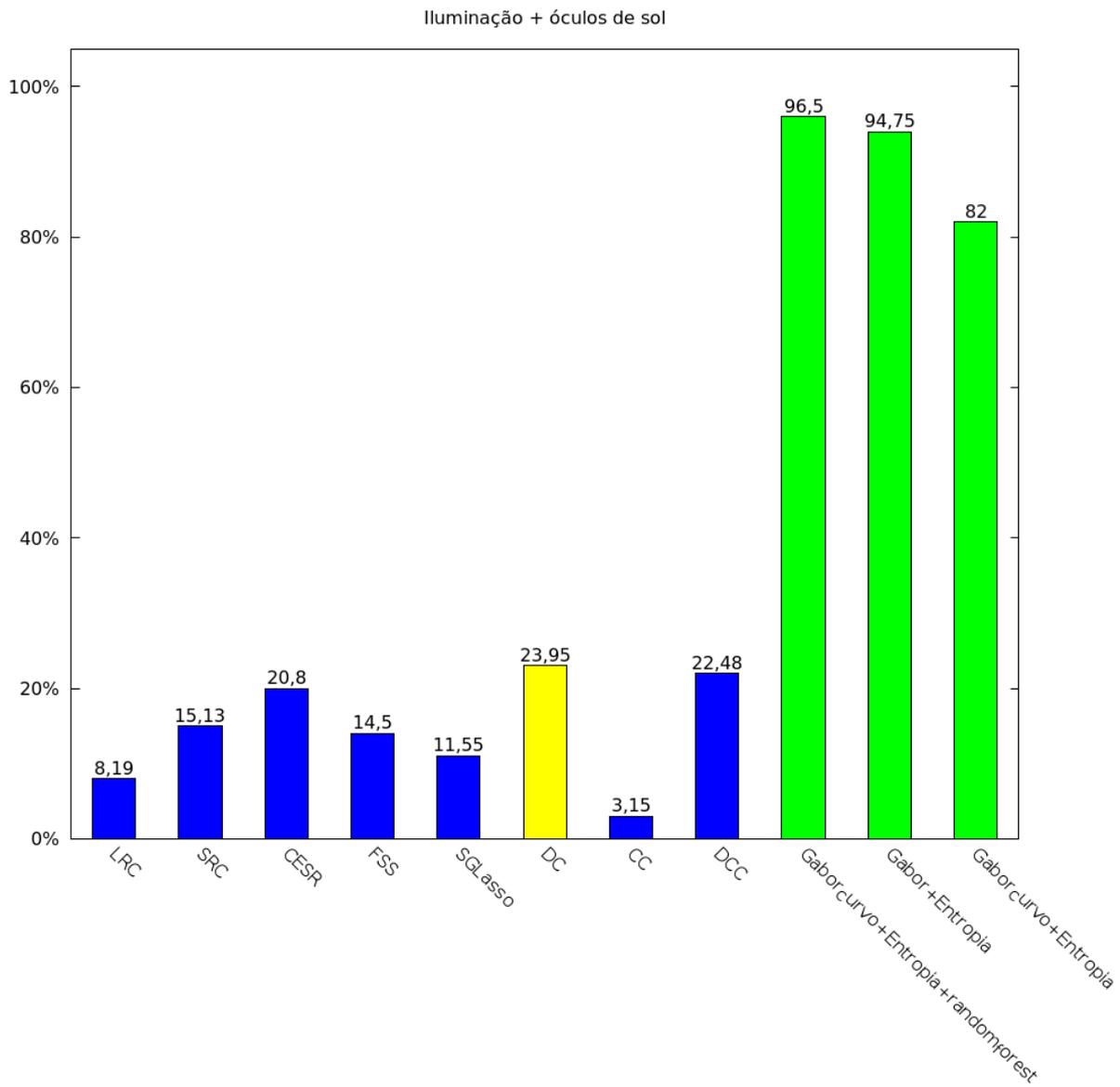


Figura 15 – Comparação de diversas abordagens com o subconjunto da *AR Face* com oclusão por óculos de sol.

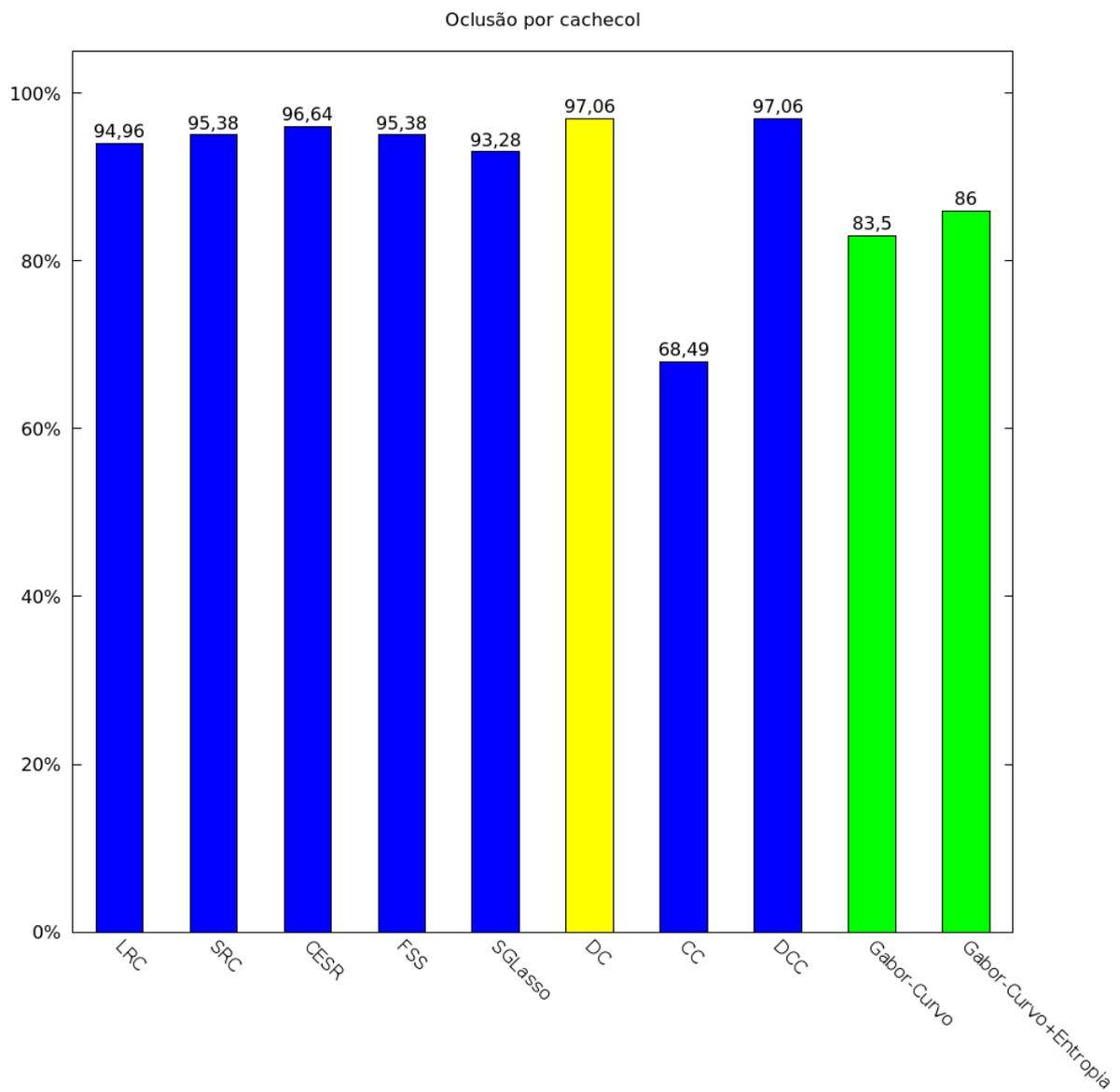


Figura 16 – Comparação de diversas abordagens com o subconjunto da *AR Face* com variação de iluminação e oclusão por óculos de sol.

Métrica	Precisão (%)	Recall (%)
Iluminação	96,81	94,33
Oclusão por óculos de sol	73,11	78,00
Iluminação + óculos de sol	93,46	92,00
Oclusão por cachecol	84,78	86,00
Iluminação + Oclusão por cachecol	97,69	97,25
Mudança de expressão facial	96,87	89,12
Base completa	98,03	98,03

Tabela 3 – Precisão e *recall* da EPGC para os subconjuntos da *AR Face*.

3.1.4 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 13d) da *AR Face*. 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 é utilizada na classificação. Nesse experimento, todos os métodos empregados obtiveram taxas de acerto elevados, sendo que 6 deles apresentaram valores superiores a 94%. As representações faciais com Gabor e entropia, no entanto produziram valores de 84,50% e 96,50% para Gabor Curvo e entropia, Figura 17. Tais métricas justificam-se pela significativa redução no tamanho da imagem final utilizada nos testes. Sugerindo, portanto, que o nível de redução de dimensionalidade aplicado às imagens deve ser ajustado com relação ao tamanho da imagem.

3.1.5 Iluminação + Oclusão por cachecol

Adicionando-se a variação de iluminação à oclusão por cachecol nas imagens 12, 13, 25 e 26 (Figura 13e), elaborou-se um cenário mais complexo para o reconhecimento facial. Para esses casos o Gabor com entropia alcança 98,25% de acerto superado pelo Gabor curvo e entropia com 99,25%, taxas acima dos 63,87% alcançados por DC, Figura 18.

3.1.6 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 13f), segrega-se 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 99,00% de acerto. Para esse cenário da *AR Face*, não foi possível efetuar comparações com as demais abordagens relacionadas nesse trabalho. As abordagens não mencionam os resultados obtidos para o subconjunto da *AR Face* com mudanças na expressão facial.

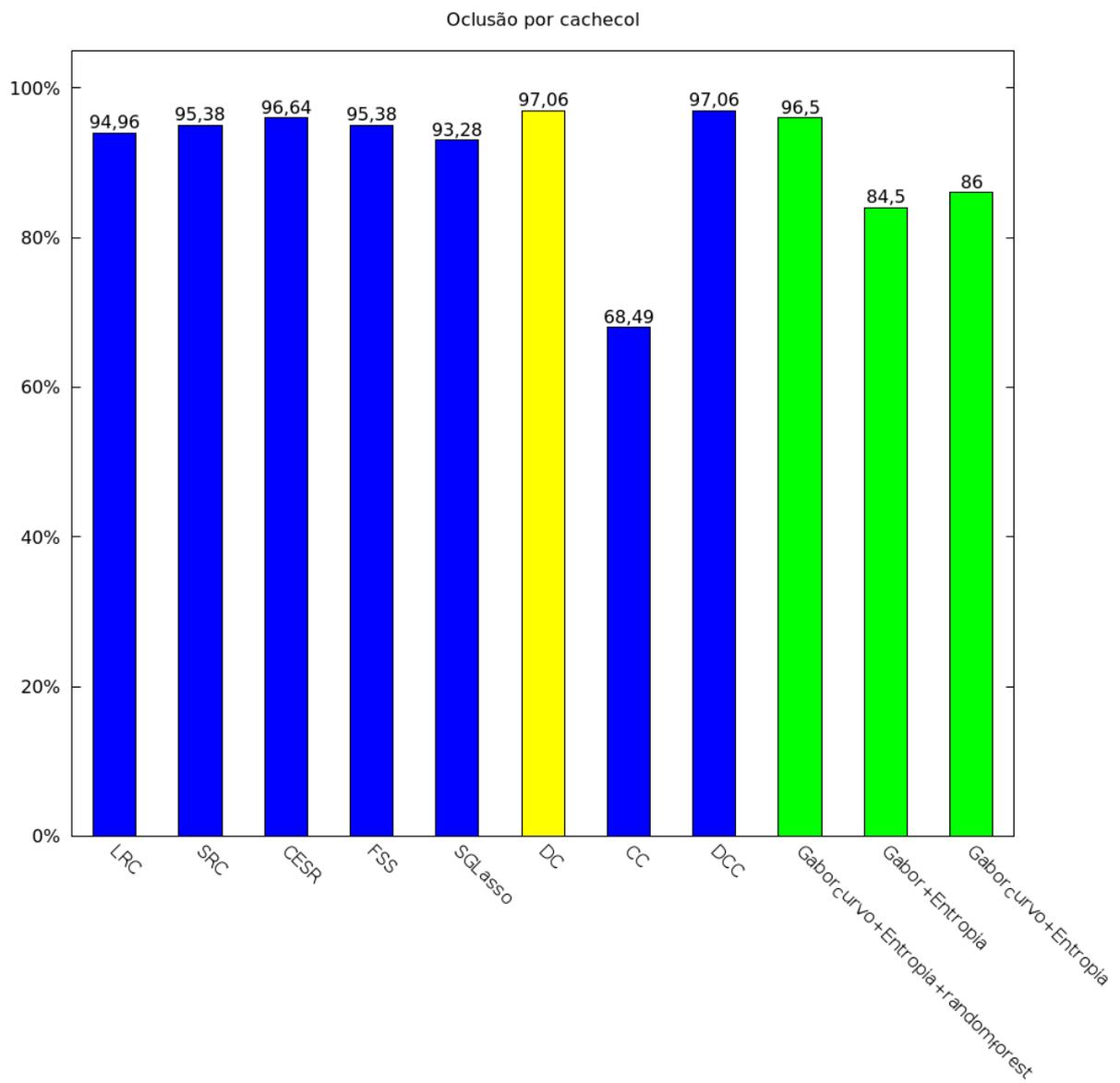


Figura 17 – Comparação de diversas abordagens com o subconjunto da *AR Face* com oclusão por cachecol.

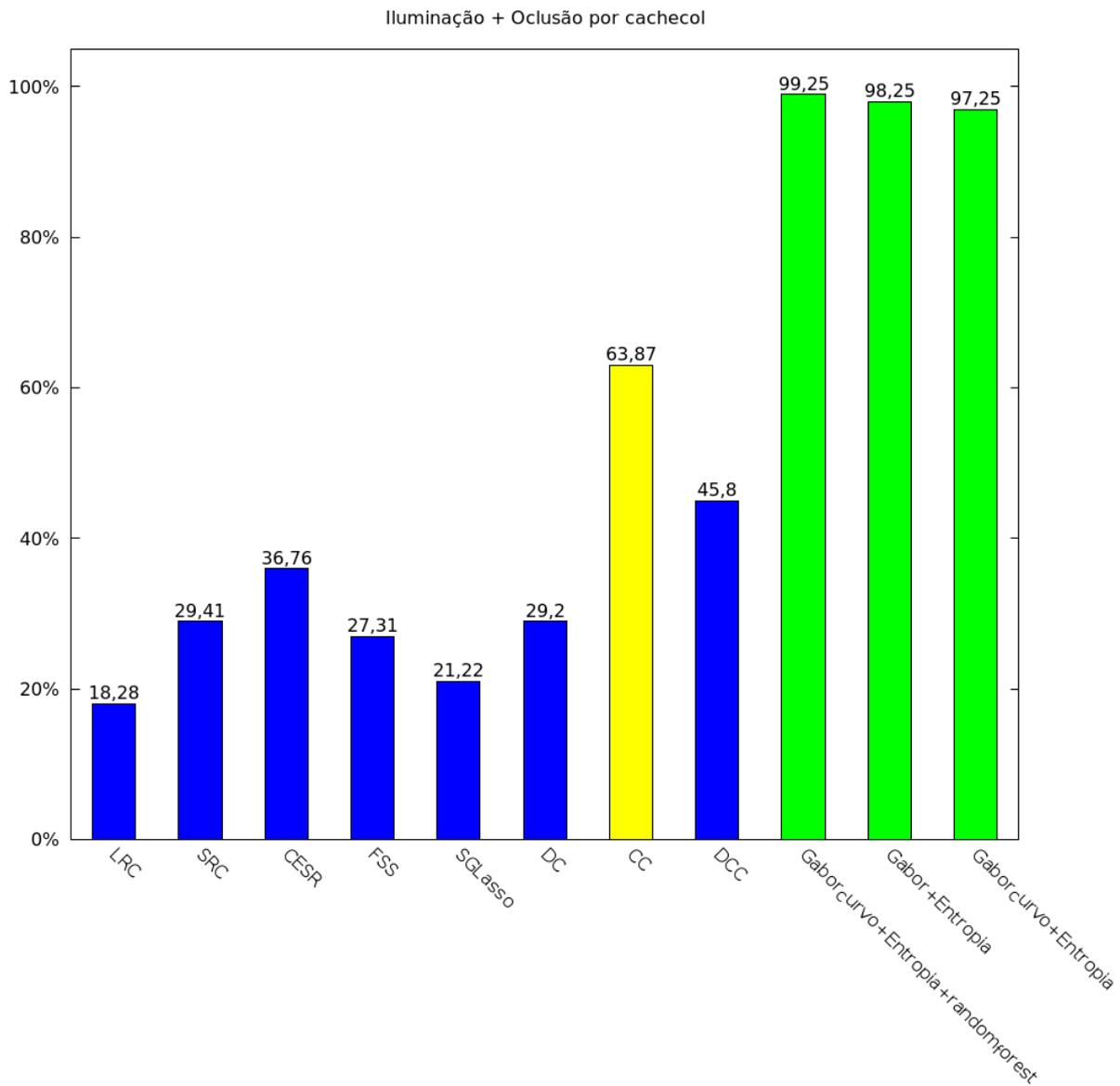


Figura 18 – Comparação de diversas abordagens com o subconjunto da *AR Face* com variação de iluminação e oclusão por cachecol.

3.1.7 Base completa

Por fim, faz-se necessária a avaliação da abordagem proposta com todas as imagens da *AR Face* (Figura 13), a fim de conseguir a simulação de um ambiente não controlado, com as possíveis ocorrências encontradas no mundo real. Nesse contexto obtém-se 99,34% de acerto para Gabor e entropia e 98,03% para Gabor Curvo e entropia com a abordagem proposta.

3.2 FERET

Constituída de 1400 imagens faciais, a base FERET possui diversas oclusões, expressões, poses e iluminações. A base de imagens foi resultado do programa de mesmo nome do Departamento de Defesa dos Estados Unidos que tinha por objetivo o desenvolvimento de um banco de imagens padrão para testes e avaliação de algoritmos de reconhecimento facial (PHILLIPS et al., 2000; YANG; DAI, 2009). No experimento apresentado, a região da face da imagem original foi manualmente cortada e as imagens redimensionadas para 48 x 56 *pixels*, equiparando-se as utilizadas nas demais abordagens de comparação. Como exemplo de imagens extraídas da FERET têm-se a Figura 19.

A Tabela 4 apresenta os resultados da acurácia das diversas abordagens na base FERET. Por meio dela verifica-se que EPGC obtém os melhores resultados com a acurácia de 97,26%, seguida de DCC com 92,84%, Figura 20. Como resultado, obtem-se uma precisão e *recall* de 97,26% indicando que a abordagem proposta tem classificação igual de indivíduos como falso positivos e falso negativos, característica apropriada para uma base balanceada.

Algoritmo	Acurácia (%)
LRC	79,22
SRC	86,37
CESR	92,35
FSS	86,27
SGLasso	82,35
DC	92,55
CC	64,02
DCC	92,84
EPGC	97,26

Tabela 4 – Comparação da acurácia de diferentes algoritmos sobre a FERET.

3.3 Yale com oclusão

Extended Yale B database corresponde a 2496 imagens faciais frontais de 38 pessoas, fotografadas em escala de cinza sobre 64 variações controladas de luminosidade. A partir

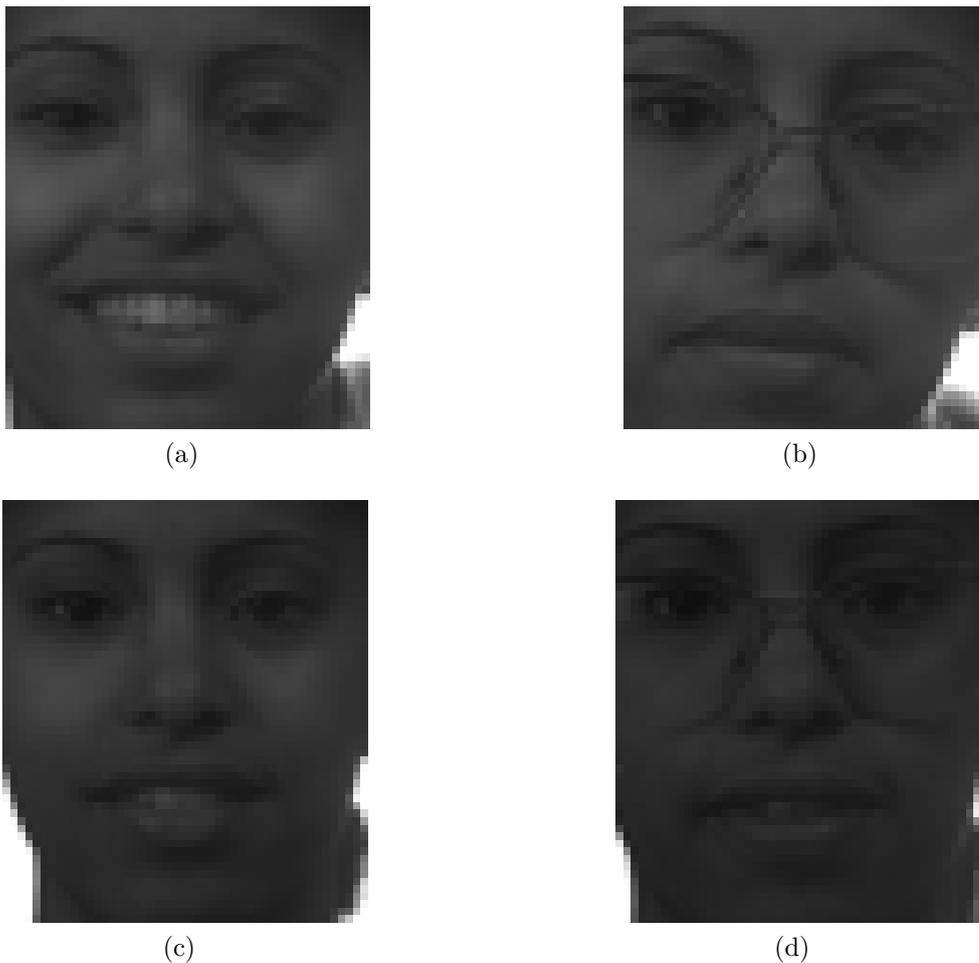


Figura 19 – Faces extraídas da *FERET*.

da base, foram escolhidas aleatoriamente 30 faces por indivíduo e sobrepostas por uma imagem fixa de um babuíno na Figura 21. O tamanho da oclusão gerada, varia entre 10%, 20%, 30%, 40%, 50%, 60%, 70% e 80% enquanto que a posição da oclusão ocorre de modo aleatório. A nova base gerada é exemplificada na Figura 22. Devido a característica aleatória do posicionamento da oclusão, as bases YALE utilizadas em cada experimento são diferentes quanto ao posicionamento da oclusão.

O resultado da classificação é mostrado na Tabela 5 e as demais métricas na Tabela 6. Esses dados apontam que em todas as abordagens comparadas, à medida que cresce a oclusão, diminui-se a acurácia. No entanto, o EPGC obtém resultados que variam de 96,22% de acurácia com 10% de oclusão à 50,87% de acurácia com 80% de oclusão. Esses valores superam as abordagens utilizadas em comparação para todas as variações de oclusão.

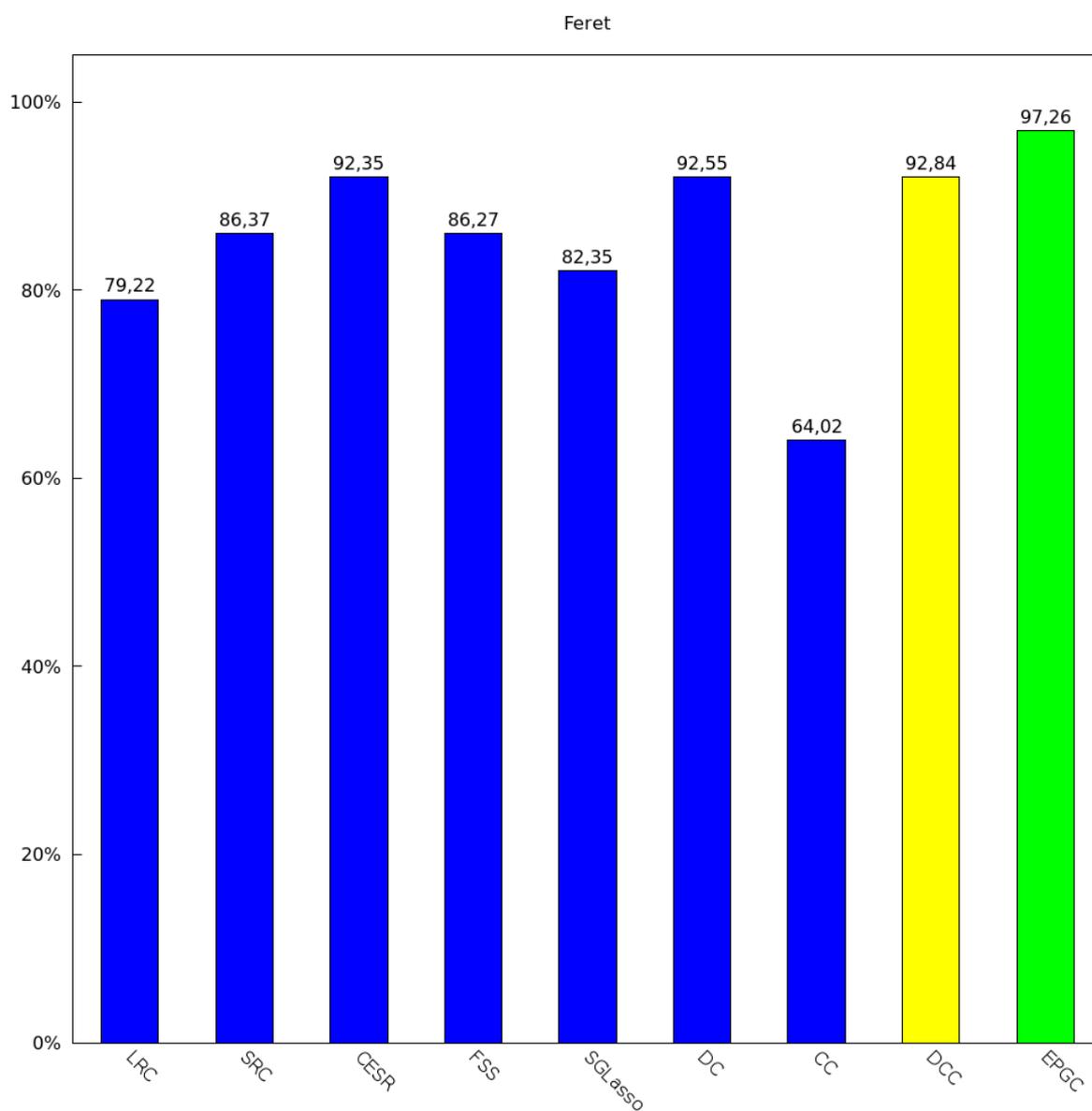


Figura 20 – Comparação de diversas abordagens com a FERET.

Oclusão	LRC	<i>SRC</i>	CESR	FSS	SGLasso	DC	CC	DCC	EPGC
10%	86,07	86,02	84,33	85,97	81,99	84,53	89,32	86,82	96,22
20%	75,16	77,56	80,90	77,47	69,88	80,31	86,55	83,19	95,43
30%	60,07	67,63	74,77	67,48	53,22	72,51	80,89	76,87	92,71
40%	49,93	56,10	69,25	56,05	42,00	67,24	74,63	71,11	90,52
50%	41,27	45,67	56,86	45,51	33,98	55,29	64,98	61,69	81,66
60%	25,95	28,11	35,00	27,97	19,54	35,49	50,53	43,59	72,36
70%	17,33	17,32	18,84	17,22	12,17	20,25	37,03	27,49	58,07
80%	11,00	9,74	10,45	9,60	9,15	11,34	23,10	17,26	50,87

Tabela 5 – Comparação dos resultados da EPGC, com abordagens sobre a base Yale com oclusão.

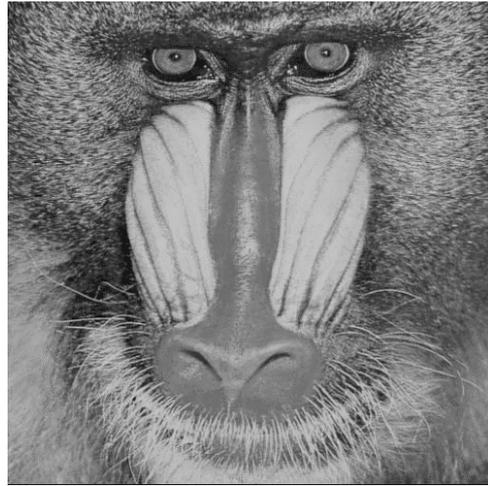


Figura 21 – Imagem de um babuíno utilizada para gerar oclusão.

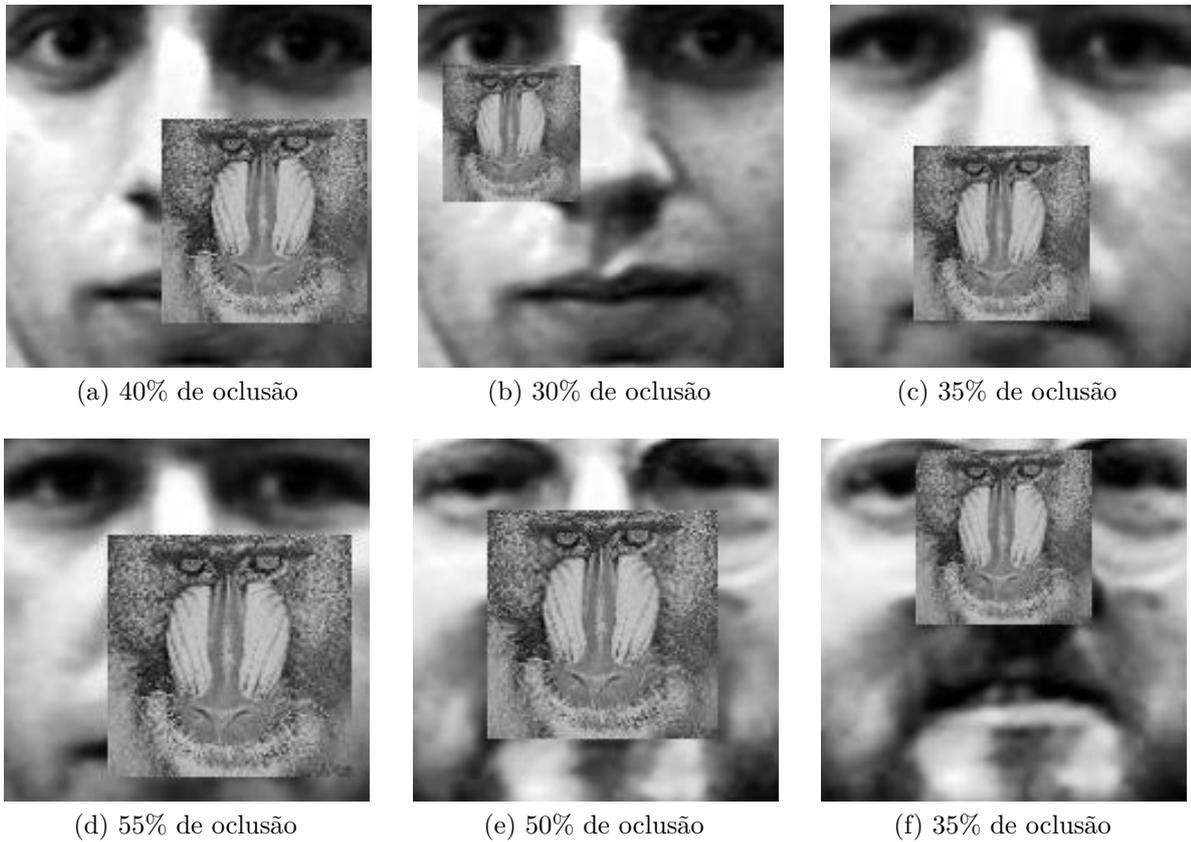


Figura 22 – Faces extraídas da Yale com oclusão.

3.4 Discussões

Os resultados obtidos pela abordagem proposta foram superiores em todos os experimentos que possuíam variação direta de iluminação, ressaltando assim a sua robustez a esse tipo de condição não controlada. Quanto aos 2 últimos testes, com mudança de expressão facial e com todas as imagens da *AR Face*, obtém-se um contexto mais próximo

Oclusão	Precisão	Recall
10%	96,22	96,22
20%	95,43	95,43
30%	92,71	92,71
40%	90,52	90,52
50%	81,66	81,66
60%	72,36	72,36
70%	58,07	58,07
80%	50,87	50,87

Tabela 6 – Precisão e *recall* referentes a Yale.

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 clássico propicia um avanço na acurácia encontrada. Tal fato pode ser justificado pela adição de mais características relevantes na representação da face, que são evidenciadas em situações com oclusão.

A abordagem EPGC possui após a etapa de extração de características (Etapa 3) um vetor de características com 94080 elementos para imagens de 48x56 *pixels*. A partir dessa representação, utilizou-se o *Random Forest* para seleção de características na Etapa 4. Como resultado, o novo vetor gerado possuiu um tamanho menor em relação ao inicial, com 2800 elementos. A redução no tamanho apresentou um percentual de 97,024%. A baixa dimensionalidade da representação da imagem propicia o uso de um menor espaço de armazenamento e tempos menores para recuperação e identificação da imagem (KAR et al., 2012).

De forma complementar, a inclusão do *randorm forest* na Etapa 3 propiciou uma evolução expressiva na taxa de acurácia. Os resultados apresentados na Tabela 2 contrapõe o uso de Gabor com Entropia, Gabor Curvo com Entropia e Gabor Curvo com Entropia e *Random Forest*, explicitando o ganho obtido entre as variações da mesma abordagem.

Na Tabela 7, têm-se o comparativo de algumas variações da abordagem proposta. Na primeira há apenas a aplicação direta do vetor de características ao SVM sem a execução da Etapa 3. Verifica-se nessa primeira abordagem valores de acurácia próximos as demais, no entanto a execução apresenta um custo elevado de uso de memória e processamento resultante do vetor de alta dimensionalidade. Nas variações da abordagem, altera-se o algoritmo utilizado para a seleção de características. Utilizou-se um segundo algoritmo, *Extremely Randomized Trees* (GEURTS; ERNST; WEHENKEL, 2006), também baseado em árvores de decisão aleatórias assim como o *Random Forest*. Verifica-se assim, que a substituição do algoritmo de seleção de características por outro da mesma categoria

mantém os resultados satisfatórios.

Os testes foram efetuados com a *AR Face*, FERET e YALE, bases muito empregadas na literatura para testes de acurácia das abordagens de reconhecimento facial. Apesar de serem bases criadas em um ambiente controlado, elas adquirem características de um ambiente não controlado, incluindo variação de iluminação, oclusão e expressões faciais, consolidando-se como bases de referência. Os experimentos de reconhecimento facial efetuados sobre as bases, possuem como etapa que requer maior poder computacional a aplicação do filtro de Gabor curvo. 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.

Base de imagem	SVM	SVM + RF + 1 Partição	SVM + RF + 2 Partições	SVM + ERT + 2 Partições
<i>AR Face</i> - Completa	98,15%	98,80%	98,03%	98,30%
Feret	94,74%	96,46%	97,26%	92,54%
Yale com oclusão de 80%	-	47,80	50,87	61,49%

Tabela 7 – Comparativo de abordagens. Na primeira não há redução do vetor de características. Na segunda e terceira abordagens é utilizado o *Random Forest*, variando o número de vezes que a Etapa 3 da EPGC é executada. Por fim, na quarta abordagem utiliza-se o *Extremely Randomized Trees* (ERT).

3.5 Considerações Finais

Neste capítulo foram apresentados os resultados relacionados ao reconhecimento facial efetuado pelo EPGC nas bases de imagens AR Face, FERET e YALE. Os resultados demonstram como a abordagem supera problemas inerentes ao reconhecimento facial, como oclusões parciais da face, variações de iluminação e expressões faciais.

4 Conclusões e Continuidade da pesquisa

Nesse trabalho, foi apresentada uma abordagem para realização de reconhecimento facial, combinando um banco de filtro de Gabor curvo, a entropia da função de projeção de variância local, *Random Forest* e SVM. Dado um conjunto de bases de imagens faciais verificou-se índices de acerto superiores, na maior parte dos casos de teste, quando comparado com algoritmos no estado da arte que indicam a eficiência da utilização do Filtro de Gabor curvo em conjunto com a entropia, *Random Forest* e SVM. Além disso, obteve-se um vetor de características compacto, representativo e robusto quando comparado com a versão inicial, a variações de iluminação (caracterizado pela aplicação do filtro Gabor curvo) e a oclusões (decorrente da segmentação em blocos).

O filtro de gabor curvo proporciona um vetor de características de tamanho elevado, sendo inviável à utilização direta por classificadores. Já o uso da entropia e *Random Forest* possibilitaram significativamente a redução do tamanho do vetor, permitindo sua classificação por meio do SVM.

A utilização de acessórios que provocam oclusões na face, como óculos de sol e cachecol, bem como diferentes expressões faciais são aspectos que afetam consideravelmente o processo de reconhecimento. Outro fator importante que interfere nos resultados dos métodos estudados é a iluminação. Como consequência do local da fonte de iluminação, diferentes sombras e valores de brilho são formados na imagem, dificultando a análise das características da face.

Por fim, vale ressaltar que a representação das imagens faciais obtém melhores resultados na classificação, 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.

4.1 Continuidade da pesquisa

Baseando-se nas limitações encontradas e em ideias para complementação deste estudo, os seguintes itens são considerados pontos primordiais para a continuidade e finalização da pesquisa:

1. Realizar experimentos em bases de dados não controladas;
2. Implementar o processo de seleção de características na Etapa 4 de forma que a execução seja distribuída, otimizando o uso de múltiplas CPUs ou GPUs.

Referências

- ALPAYDIN, E. Machine learning, statistics, and data analytics. In: _____. *Machine Learning: The New AI*. MIT Press, 2016. p. 224–. ISBN 9780262337595. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7845219>>. Citado na página 3.
- BELLMAN, R. *Adaptive Control Processes*. [S.l.]: Princeton University Press, 1961. Citado na página 3.
- BISHOP, C. M. *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press, 1995. Citado na página 3.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001. Citado na página 13.
- CAMENT, L. A. et al. Face recognition under pose variation with active shape model to adjust gabor filter kernels and to correct feature extraction location. In: IEEE. *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. [S.l.], 2015. v. 1, p. 1–6. Citado na página 10.
- CHAI, Z. et al. Gabor ordinal measures for face recognition. *IEEE transactions on information forensics and security*, IEEE, v. 9, n. 1, p. 14–26, 2014. Citado na página 9.
- CHAN, T.-H. et al. Pcanet: A simple deep learning baseline for image classification? *IEEE Transactions on Image Processing*, IEEE, v. 24, n. 12, p. 5017–5032, 2015. Citado na página 21.
- CHANG, C.-C.; LIN, C.-J. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, ACM, New York, NY, USA, v. 2, n. 3, p. 27:1–27:27, maio 2011. ISSN 2157-6904. Citado 2 vezes nas páginas 17 e 24.
- CHANG, C.-I. et al. Survey and comparative analysis of entropy and relative entropy thresholding techniques. *IEE Proceedings - Vision, Image and Signal Processing*, v. 153, n. 6, p. 837–850, Dec 2006. ISSN 1350-245X. Citado na página 15.
- CHELLAPPA, R.; NI, J.; PATEL, V. M. Remote identification of faces: Problems, prospects, and progress. *Pattern Recognition Letters*, v. 33, n. 14, p. 1849–1859, 2012. Citado na página 1.
- CHEN, W.-S. et al. A novel discriminant criterion based on feature fusion strategy for face recognition. *Neurocomputing*, Elsevier, v. 159, p. 67–77, 2015. Citado na página 2.
- CORTES, C.; VAPNIK, V. Support Vector Networks. *Machine Learning*, v. 20, p. 273–297, 1995. Citado na página 17.
- DAUGMAN, J. G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A*, Optical Society of America, v. 2, n. 7, p. 1160–1169, 1985. Citado na página 9.

FENG, G.-C.; YUEN, P. C. Variance projection function and its application to eye detection for human face recognition. *Pattern Recognition Letters*, Elsevier, v. 19, n. 9, p. 899–906, 1998. Citado 2 vezes nas páginas 14 e 22.

GABOR, D. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, IET, v. 93, n. 26, p. 429–441, 1946. Citado na página 9.

GEORGHIADES, A.; BELHUMEUR, P.; KRIEGMAN, D. From Few to Many: Illumination dels for Face Recognition under Variable Lighting and Pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, v. 23, n. 6, p. 643–660, 2001. Citado na página 29.

GEURTS, P.; ERNST, D.; WEHENKEL, L. Extremely randomized trees. *Machine learning*, Springer, v. 63, n. 1, p. 3–42, 2006. Citado na página 43.

HAGHIGHAT, M. B. A.; NAMJOO, E. Evaluating the informativity of features in dimensionality reduction methods. In: IEEE. *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*. [S.l.], 2011. p. 1–5. Citado 2 vezes nas páginas 3 e 19.

HE, R.; ZHENG, W. S.; HU, B. G. Maximum Correntropy Criterion for Robust Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 33, n. 8, p. 1561–1576, Aug 2011. ISSN 0162-8828. Citado 2 vezes nas páginas 5 e 29.

HEISELE, B.; HO, P.; POGGIO, T. Face recognition with support vector machines: Global versus component-based approach. In: IEEE. *Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001*. [S.l.], 2001. v. 2, p. 688–694. Citado na página 16.

HWANG, W. et al. Face recognition using Extended Curvature Gabor classifier bunch. *Pattern Recognition*, v. 48, n. 4, p. 1247–1260, 2015. Citado 3 vezes nas páginas 3, 12 e 19.

JANITZA, S.; TUTZ, G.; BOULESTEIX, A.-L. Random forest for ordinal responses: Prediction and variable selection. *Computational Statistics & Data Analysis*, Elsevier, v. 96, p. 57–73, 2016. Citado na página 24.

KAMARAINEN, J.-K.; KYRKI, V.; KÄLVIÄINEN, H. Invariance properties of gabor filter-based features-overview and applications. *IEEE Transactions on Image Processing*, IEEE, v. 15, n. 5, p. 1088–1099, 2006. Citado na página 9.

KAR, A. et al. Human face recognition using gabor based kernel entropy component analysis. *International Journal of Computer Vision and Image Processing (IJCVIP)*, IGI Global, v. 2, n. 3, p. 1–20, 2012. Citado 3 vezes nas páginas 3, 4 e 43.

KNERR, S.; PERSONNAZ, L.; DREYFUS, G. Single-Layer Learning Revisited: A Stepwise Procedure for Building and Training a Neural Network. In: Fogelman Soulié, F.; HÉRAULT, J. (Ed.). *Neurocomputing: Algorithms, Architectures and Applications*. [S.l.]: Springer-Verlag, 1990, (NATO ASI Series, F68). p. 41–50. Citado na página 17.

KRESSEL, U. H.-G. Pairwise classification and support vector machines. In: MIT PRESS. *Advances in kernel methods*. [S.l.], 1999. p. 255–268. Citado na página 17.

- LAI, Z. R. et al. Discriminative and Compact Coding for Robust Face Recognition. *IEEE Transactions on Cybernetics*, v. 45, n. 9, p. 1900–1912, Sept 2015. ISSN 2168-2267. Citado 3 vezes nas páginas 5, 6 e 29.
- LEE, H. et al. Efficient sparse coding algorithms. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2006. p. 801–808. Citado 3 vezes nas páginas 5, 6 e 29.
- LI, S. Z.; JAIN, A. K. *Handbook of Face Recognition*. [S.l.]: Springer Publishing Company, 2011. Citado 4 vezes nas páginas 15, 1, 2 e 3.
- LIAO, P. et al. Parameter optimization for support vector machine based on nested genetic algorithms. *Journal of Automation and Control Engineering*, v. 3, n. 6, 2015. Citado na página 25.
- LIU, C.; WECHSLER, H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image processing*, IEEE, v. 11, n. 4, p. 467–476, 2002. Citado na página 10.
- LIU, C.; WECHSLER, H. Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition. *Transactions on Image Processing*, IEEE Press, Piscataway, NJ, USA, v. 11, n. 4, p. 467–476, abr. 2002. ISSN 1057-7149. Citado 2 vezes nas páginas 11 e 19.
- MARTINEZ, A. M. The AR face database. *CVC Technical Report*, v. 24, 1998. Citado na página 29.
- MESHGINI, S.; AGHAGOLZADEH, A.; SEYEDARABI, H. Face recognition using Gabor-based direct linear discriminant analysis and support vector machine. *Computers & Electrical Engineering*, v. 39, n. 3, p. 727–745, 2013. Citado na página 9.
- NASEEM, I.; TOGNERI, R.; BENNAMOUN, M. Linear Regression for Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 32, n. 11, p. 2106–2112, Nov 2010. ISSN 0162-8828. Citado 2 vezes nas páginas 5 e 29.
- PANG, W.-M.; CHOI, K.-S.; QIN, J. Fast gabor texture feature extraction with separable filters using gpu. *Journal of Real-Time Image Processing*, v. 12, n. 1, p. 5–13, 2016. ISSN 1861-8219. Disponível em: <<http://dx.doi.org/10.1007/s11554-013-0373-y>>. Citado na página 21.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, v. 12, n. Oct, p. 2825–2830, 2011. Citado na página 29.
- PHILLIPS, P. J. et al. The FERET Evaluation Methodology for Face-Recognition Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Washington, DC, USA, v. 22, n. 10, p. 1090–1104, out. 2000. ISSN 0162-8828. Citado 3 vezes nas páginas 1, 29 e 39.
- PIEPERS, D. W.; ROBBINS, R. A. A review and clarification of the terms “holistic,” “configural,” and “relational” in the face perception literature. *Frontiers in psychology*, Frontiers Media SA, v. 3, 2012. Citado na página 6.
- REN, C.-X. et al. Band-reweighed gabor kernel embedding for face image representation and recognition. *IEEE Transactions on Image Processing*, IEEE, v. 23, n. 2, p. 725–740, 2014. Citado 2 vezes nas páginas 9 e 11.

- SCHÖLKOPF, B.; BURGESS, C.; VAPNIK, V. Extracting support data for a given task. In: *First International Conference on Knowledge Discovery and Data Mining*. Menlo Park: AAAI Press, 1995. Citado na página 17.
- SHANNON, C. A mathematical theory of communication. *Bell System Technical Journal, The*, v. 27, n. 3, p. 379–423, July 1948. ISSN 0005-8580. Citado na página 15.
- SHEN, L.; BAI, L.; FAIRHURST, M. C. Gabor wavelets and General Discriminant Analysis for face identification and verification. *Image and Vision Computing*, v. 25, n. 5, p. 553–563, 2007. Citado 3 vezes nas páginas 9, 10 e 19.
- SRINIVAS, U. et al. Structured sparse priors for image classification. *IEEE Transactions on Image Processing*, IEEE, v. 24, n. 6, p. 1763–1776, 2015. Citado na página 1.
- THE MATHWORKS, INC. *MATLAB version 8.5.0.197613 (R2014b)*. Natick, Massachusetts, 2014. Citado na página 21.
- TORKHANI, G. et al. A 3d–2d face recognition method based on extended gabor wavelet combining curvature and edge detection. *Signal, Image and Video Processing*, Springer, p. 1–8, 2017. Citado na página 11.
- VALIANT, L. A Theory of the Learnable. *Communications of the ACM*, v. 27, n. 11, p. 1134–1142, 1984. Citado na página 3.
- VAPNIK, V. N. *The Nature of Statistical Learning Theory*. [S.l.]: Springer, 1995. Citado 2 vezes nas páginas 5 e 15.
- WENG, R.; LU, J.; TAN, Y. P. Robust point set matching for partial face recognition. *IEEE Transactions on Image Processing*, v. 25, n. 3, p. 1163–1176, March 2016. ISSN 1057-7149. Citado na página 6.
- WISKOTT, L. et al. Face recognition by elastic bunch graph matching. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 19, n. 7, p. 775–779, 1997. Citado na página 10.
- WITTEN, I. H. et al. *Data Mining: Practical machine learning tools and techniques*. [S.l.]: Morgan Kaufmann, 2016. Citado na página 3.
- WONG, Y. et al. On robust biometric identity verification via sparse encoding of faces: Holistic vs local approaches. In: IEEE. *The 2012 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2012. p. 1–8. Citado na página 6.
- WRIGHT, J. et al. Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 31, n. 2, p. 210–227, Feb 2009. ISSN 0162-8828. Citado 3 vezes nas páginas 5, 6 e 29.
- YANG, W.-H.; DAI, D.-Q. Two-dimensional maximum margin feature extraction for face recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, IEEE, v. 39, n. 4, p. 1002–1012, 2009. Citado na página 39.
- YUAN, M.; LIN, Y. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 68, n. 1, p. 49–67, 2006. Citado 3 vezes nas páginas 5, 6 e 29.

ZHANG, J.-S.; CHEN, C.-J. Local variance projection log energy entropy features for illumination robust face recognition. In: IEEE. *International Symposium on Biometrics and Security Technologies, 2008. ISBAST 2008*. [S.l.], 2008. p. 1–5. Citado 2 vezes nas páginas [21](#) e [22](#).

ZHOU, Z.-H.; GENG, X. Projection functions for eye detection. *Pattern recognition*, Elsevier, v. 37, n. 5, p. 1049–1056, 2004. Citado 2 vezes nas páginas [14](#) e [22](#).