Um estudo comparativo das estratégias de fusão no nível de característica para Sistemas Biométricos Multimodais baseados em face e íris

Alternative Title: A comparative study of feature level fusion strategies for Multimodal Biometric Systems based on Face and Iris

Daniel M. M da Costa
Universidade de São Paulo
São Paulo - SP - Brasil
daniel.moura.costa@usp.br

Henrique Passos
Universidade de São Paulo
São Paulo - SP - Brasil
henrique.passos@usp.br

Sarajane Marques Peres
Universidade de São Paulo
São Paulo - SP - Brasil
sarajane@usp.br

Clodoaldo A. M. de Lima
Universidade de São Paulo
São Paulo - SP - Brasil
c.lima@usp.br

RESUMO

Com o avanço da tecnologia, novas abordagens para o reconhecimento automático da identidade de uma pessoa têm sido propostas e tal fato tem encorajado o emprego de Sistemas Biométricos. Essa abordagem utiliza características físicas ou comportamentais de uma pessoa para realizar a sua identificação. Os Sistemas Biométricos podem ser classificados como Unimodais ou Multimodais. Sistemas Biométricos Unimodais utilizam apenas uma modalidade biométrica para realizar o reconhecimento, ao passo que os Sistemas Biométricos Multimodais empregam duas ou mais modalidades. A construção de um Sistema Biométrico Multimodal pode ser realizada de diferentes formas, as quais são categorizadas de acordo com a sua arquitetura, nível de fusão e estratégia de fusão. O objetivo deste trabalho é investigar diferentes formas de fusão de modalidades biométricas no nível de característica, visando projetar um sistema multimodal com alto poder de reconhecimento. Neste artigo, nós utilizamos a Transformada Wavelet para extrair um conjunto de características de imagens de face e íris. Os resultados obtidos mostram que os Sistemas Biométricos Multimodais apresentam melhor desempenho que os Sistemas Biométricos Unimodais, em termos de taxa de reconhecimento calculada sobre a saída produzida pela Máquina de Vetores Suporte usada como classificador.

Palavras-Chave

Biometria Multimodal, Biometria Unimodal, Transformada Wavelet, Máquina de Vetores Suporte.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
SBSI 2015, May 26th-29th, 2015, Goiânia, Goiás, Brazil
Copyright SBC 2015.

ABSTRACT

With the technology advances, new approaches for automatic recognition of a person's identity have been proposed and such a fact has encouraged the use of Biometrics Systems. This approach uses physical or behavioural characteristics of the user in order to recognize or authenticate their identity. The Biometric Systems can be classified as Unimodal or Multimodal. The Unimodal Systems use a single biometric modality to perform the recognition, while the Multimodal ones use two or more modalities. A Multimodal Biometric System can be constructed in different ways, according to its architecture, fusion level and fusion strategies. The main of this work is to investigate and compare different feature level fusion strategies, in order to design a Multimodal Biometric System with high performance. In this paper, we used the discrete wavelet transform to extract the feature sets from iris and face images. Experimental results show that Multimodal Biometric Systems outperform Unimodal Biometric Systems according to recognition rate computed over the outputs produced by the induced Support Vector Machine classifier.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Miscellaneous; D.2.8 [**Pattern Recognition**]: Application

General Terms

Design, Theory

Keywords

Multimodal Biometrics, Unimodal Biometrics, Wavelet Transform, Support Vector Machines.

1. INTRODUÇÃO

Devido à necessidade de fornecer segurança sobre a informação gerada em diferentes contextos, como por exemplo, numa transação monetária realizada pela internet ou

no controle de acesso a determinados locais, geralmente, são utilizadas estratégias como senhas, tokens, números PIN ou documentos pessoais para identificar o usuário e confirmar o seu acesso. Apesar dos métodos tradicionais obterem um bom desempenho na prática, estes estão sujeitos à fraude e à falsificação, visto que as informações podem ser roubadas, descobertas ou falsificadas por outros usuários [11].

Neste contexto, existe um crescente incentivo ao uso da tecnologia biométrica para melhorar, e até mesmo substituir, os métodos convencionais de segurança. A biometria emprega características físicas ou comportamentais do usuário para confirmar a sua identidade [10]. Uma das vantagens do uso dessa abordagem é que o usuário precisa interagir diretamente com o sistema, tornando a sua invasão mais difícil. As principais modalidades que podem ser usadas como biometria são a impressão digital [1], íris [24], face [15] e recentemente os sinais cerebrais [22] e cardíacos [16].

Para que uma determinada característica pessoal possa ser considerada uma modalidade biométrica, algumas propriedades devem ser satisfeitas [10] [9]: i) Universalidade: todas as pessoas devem possuir essa característica; ii) Distinção: por meio dessa característica deve ser possível distinguir duas pessoas quaisquer; iii) Mensurabilidade: deve permitir uma análise quantitativa da informação; iv) Desempenho: o reconhecimento deve ser rápido e com baixa taxa de erro; v) Não invasiva: não deve causar desconforto ao usuário durante a captura dessa característica.

Dependendo do contexto, um sistema biométrico pode atuar de duas formas diferentes [10], como um sistema de verificação (autenticação) realizando uma comparação um para um, cujo objetivo é confirmar ou negar a identidade reivindicada pelo indivíduo. Neste caso o sistema visa responder a seguinte questão: "Esta pessoa é quem ela diz ser?". Já na identificação (reconhecimento), o sistema realiza uma comparação de um para muitos, o qual visa encontrar o cadastro pertencente ao indivíduo em um banco de dados. Neste o sistema é projetado para responder a questão: "Quem é esta pessoa?". Para tornar possível o uso do sistema é necessário que o indivíduo primeiramente faça seu cadastro para que em um segundo momento possa ser realizado a requisição de autenticação ou identificação. Este artigo irá focar principalmente na identificação.

Um sistema biométrico é composto de quatro módulos principais [10], como visto na Figura 1: módulo sensorial, responsável por capturar os dados do usuário; módulo de extração de características, extrai a informação útil do dado capturado; módulo de comparação, compara a informação extraída com as armazenadas no banco de dados; módulo de decisão, recebe a saída do módulo anterior e decide quem é o usuário ou se o usuário será aceito ou rejeitado.

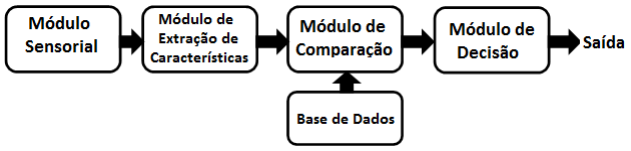

Figura 1: Arquitetura de um sistema biométrico.

Nos últimos anos, o reconhecimento biométrico tem alcançado melhorias em termos de confiabilidade e precisão, com algumas modalidades biométricas oferecendo um bom desempenho global. No entanto, mesmos os sistemas mais avançados ainda enfrentam muitos problemas, tais como [19]: i) Dados de entrada ruidosos; ii) Variações intraclasse; iii) Similaridade interclasses; iv) Não-universalidade; v) Possibilidade de falsificação;

Visando atacar estes problemas, algumas abordagens têm sido proposta na literatura. Dentre estas abordagens destaca-se a Biometria Multimodal, que emprega diferentes modalidades biométricas em conjunto, como forma de gerar sistemas capazes de superar as dificuldades enfrentadas pelos Sistemas Unimodais. Biometria multimodal pode melhorar significativamente o desempenho de reconhecimento, além de aumentar a cobertura da população, impedindo ataques por repetição e reduzindo a taxa de falsa aceitação.

Para a construção de um Sistema Multimodal é realizada a união de dois, ou mais, sistemas biométricos unimodais. Essa união pode ser realizada de diferentes maneiras e são categorizadas de acordo com a sua arquitetura, nível de fusão e estratégia de fusão [10]. Em relação ao nível da fusão, as modalidades biométricas podem ser integradas no nível sensorial, extração de característica, comparação e decisão. De forma geral, a chave para o sucesso de um sistema biométrico multimodal está ligada a um esquema efetivo de fusão. Este trabalho irá investigar formas de integração no nível de característica de duas modalidades biométricas: face e íris.

Um dos principais módulos em um sistema biométrico é o módulo de extração de características. Neste módulo, após a etapa de pré-processamento das imagens de entrada, seja esta contendo a face ou a íris de um usuário, pode-se usar os pixels como entrada para um classificador. A vantagem de utilizarmos características extraídas das imagens, ao invés dos pixels, é que estes geram uma melhor separação interclasses e uma menor variação intraclasse representando melhor o domínio da aplicação. Dentre as diversas técnicas de extração de característica destaca-se a Transformada Wavelet (WT). Esta gera um conjunto de características, chamadas de coeficientes de Wavelet, que podem ser utilizados para representar a imagem original. Neste artigo, usaremos WT como extrator de característica.

De forma a projetar um sistema biométrico, unimodal ou multimodal, é necessário definir um classificador que será responsável por receber as características extraídas do usuário e verificar se o usuário tem, ou não, permissão para acessar um determinado recurso. Dentre os possíveis classificadores que podem ser utilizadas temos: Redes Neurais Artificiais [20], Máquinas de Vetores Suporte (SVM) [21] e K-Vizinhos Mais Próximos [12]. Aqui usaremos Máquinas de Vetores Suporte para compor o módulo de comparação do nosso sistema. A SVM [27] é um classificador binário que tem sido utilizado com sucesso em diversos problemas de classificação. De forma a garantir um bom desempenho ao usar esse classificador, alguns parâmetros devem ser escolhidos cuidadosamente, como a função *Kernel* a ser utilizada, seus parâmetros e o parâmetro de regularização.

O objetivo deste trabalho foi investigar diferentes formas de combinação de modalidades biométricas no nível de característica, visando projetar um sistema biométrico multimodal de alto desempenho. Duas modalidades biométricas foram utilizadas e como técnica de extração de característica foi empregada Transformada Wavelet, sendo adotada no nível de comparação às Máquinas de Vetores Suporte.

Esse trabalho está organizado nas seguintes seções: Seções 2 e 3 apresentam os conceitos e os métodos para construção de sistemas biométricos unimodais e multimodais, nas Se-

ções 4 e 5 são apresentadas uma introdução sobre Transformada Wavelet e Máquinas de Vetores Suporte, sendo essas utilizadas nos módulos de extração de características e de classificação, respectivamente. A Seção 6 descreve os procedimentos adotados para a realização dos experimentos, Seção 7 apresenta o desempenho obtido com os sistemas biométricos construídos. Finalmente, na Seção 8 é apresentado uma breve conclusão sobre os resultados e os próximos passos a serem adotados.

2. SISTEMAS BIOMÉTRICOS UNIMODAIS

Os Sistemas Biométricos Unimodais foram os primeiros a serem propostos na literatura. Estes empregam apenas uma modalidade biométrica para realizar o reconhecimento do usuário. Um Sistema Biométrico Unimodal baseado em face pode ser descrito como um sistema que usa uma câmera fotográfica para capturar uma imagem contendo a face do usuário e então verifica se o usuário está ou não cadastrado no sistema. Este é considerado um sistema de baixo custo de construção, já que qualquer câmera pode ser utilizada para capturar a face, porém quanto maior a resolução da imagem, mais informação pode ser extraída, facilitando o reconhecimento do usuário[6].

Apesar da imagem de capturada poder ser utilizada diretamente no sistema, geralmente, os sistemas empregam algoritmos de detecção de face para reduzir a interferência do fundo da imagem no desempenho do classificador. Diversos métodos têm sido propostos para detecção de face, dentre estes, podemos destacar o Algoritmo de Viola-Jones [29]. Este algoritmo emprega um conjunto de classificadores organizados em cascata para localizar a face em uma imagem. Neste trabalho foi utilizado esse algoritmo para destacar as faces e remover o fundo da imagem. Na Figura 2 temos um exemplo das faces detectadas e segmentadas.

Figura 2: Detecção de face usando o algoritmo de Viola-Jones.

Um dos grandes desafios enfrentados pelos Sistemas Biométricos baseados em Face está na detecção da face do usuário em ambientes com uma grande variação de iluminação, poses ou com a presença de acessórios [6]. Devido a essas limitações, os Sistemas Biométricos baseados em face são geralmente utilizados em ambientes controlados, que são locais onde esses algoritmos funcionam perfeitamente.

Já os Sistemas Biométricos baseados em íris empregam sensores específicos, ou câmeras de alta-resolução, para fazer a captura da íris do usuário. Apesar da íris não ser amplamente utilizada, essa é considerada uma modalidade biométrica de alto desempenho e confiabilidade, visto que o seu processo de geração é totalmente aleatório [23] e que a sua morfologia não sofre alteração, com exceção em casos específicos de algumas doenças [8].

Da mesma forma que os sistemas faciais, os sistemas envolvendo a íris, geralmente, empregam algoritmos de detecção e segmentação da íris de forma a separar a esclera e a pupila da imagem. Neste caso, é utilizada apenas a região da íris para fazer o reconhecimento do usuário.

Para a detecção da íris utilizamos o Método de Wildes e Camus [3], que retorna um conjunto de pontos (x, y, r_p, r_i), onde x e y são as coordenadas do centro da pupila, r_p é o raio da pupila e r_i o raio da íris. Utilizando esse conjunto de pontos e o Modelo de Daugman [5], podemos realizar a segmentação da íris em coordenadas polares. Sobre essa nova imagem é realizada o processo de extração de características. Um exemplo desse processo pode ser visto na Figura 3.

Figura 3: Processo de detecção e segmentação da Íris.

3. SISTEMAS BIOMÉTRICOS MULTIMODAIS

De forma a superar as limitações enfrentadas pelos Sistemas Biométricos Unimodais, foram propostos os chamados Sistemas Biométricos Multimodais. Estes sistemas fazem o uso de duas, ou mais, fontes de informação biométrica.

De forma geral a área de Biometria Multimodal apresenta as seguintes vantagens em relação às abordagens unimodais: i) melhora significativa no desempenho de reconhecimento - diferentes modalidades medem informações complementares e por isto podem alcançar melhor desempenho; ii) melhora na robustez do sistema frente a variações do ambiente - todas as modalidades biométricas sofrem com algum tipo de interferência do ambiente, logo ao realizarmos a fusão de duas modalidades, espera-se uma redução no nível de interferência do ambiente; iii) diminui o problema da não-universalidade - ao utilizarmos mais de uma modalidade biométrica existe uma maior probabilidade de cobertura da população; iv) gera uma boa defesa contra ataques por circunvenção e repetição - o grau de dificuldade para que mais de uma modalidade biométrica seja falsificada ao mesmo tempo é maior, além disso essas características podem ser captadas de forma aleatória.

Para a construção de um Sistema Biométrico Multimodal é necessário que algumas decisões sejam tomadas, como por exemplo, quais modalidades serão utilizadas, quais características serão extraídas de cada modalidade, em qual nível as características serão combinadas e qual será o método de fusão. Em [25] são apresentadas diferentes formas de construção de sistemas biométricos multimodais.

Após a extração das características de cada modalidade biométrica, é necessário que seja definido em qual momento será realizada a fusão e qual estratégia que será utilizada para fazer essa fusão. A Figura 4 apresenta as principais formas de fusão comumente empregadas. Em [25] essas estratégias são descritas detalhadamente. Neste artigo, iremos focar na fusão no nível de característica.

3.1 Estratégias para Combinação de Características

Em [7] foram propostas algumas estratégias de fusão, no nível de característica, de diferentes modalidades biométricas usando como técnica de extração de característica a

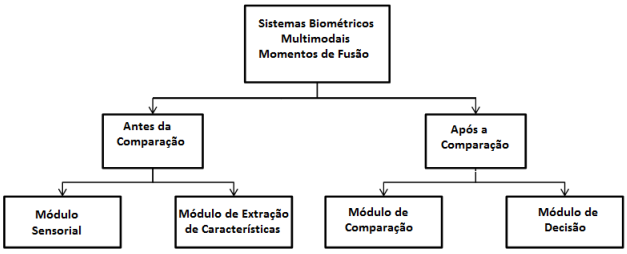

Figura 4: Níveis de fusão das características biométricas.

Transformada Wavelet. Basicamente, as estratégias propostas faziam a substituição de alguns dos coeficientes de Wavelet de uma determinada modalidade biométrica, usando alguma medida heurística, pelos coeficientes de Wavelet de outras modalidades. Com base nos coeficientes resultantes, uma imagem era construída e utilizada como entrada para o classificador. O problema desta estratégia é que a imagem reconstruída tinha a mesma dimensão da imagem original, resultando num conjunto de atributos de alta dimensionalidade. Visando contornar este problema, as estratégias de fusão adotadas neste artigo é realizada no nível de decomposição igual a 3, resultando em uma quantidade menor de coeficientes, sendo esses coeficientes utilizados como entrada para o classificador. As seguintes estratégias de fusão foram adotadas:

- *Estratégia #1* - nessa estratégia selecionamos os coeficientes de aproximação da face e concatenamos com os coeficientes de detalhes da íris;
- *Estratégia #2* - nessa estratégia selecionamos os coeficientes de aproximação da íris e concatenamos com os coeficientes de detalhes da face;
- *Estratégia #3* - nessa estratégia calculamos a média dos coeficientes de aproximação e concatenamos com máximo dos coeficientes de detalhes entre a face e a íris;
- *Estratégia #4* - nessa estratégia selecionamos o mínimo dos coeficientes de aproximação entre a face e a íris e concatenamos com máximo dos coeficientes de detalhes entre a face e a íris;

Após a aplicação das técnicas de fusão foram gerados novos vetores de características, sendo esses utilizados para representar a fusão da face e a íris do usuário.

4. TRANSFORMADA WAVELET

A Transformada de Fourier provê uma boa descrição das frequências de um sinal, mas não do seu comportamento temporal. Uma das primeiras abordagens para suprir essa carência da transformada de Fourier (e oferecer capacidade de análise no tempo e na frequência) foi subdividir um sinal em segmentos e calcular a transformada de Fourier em cada um deles. Essa abordagem é conhecida como Transformada de Fourier de Curto Tempo. A Transformada de Fourier de Curto Tempo utiliza uma função para segmentar o sinal original (criar janelas), e então aplica a transformada de Fourier em cada um dos segmentos gerados. Este método tem sido utilizado com sucesso em aplicações biomédicas.

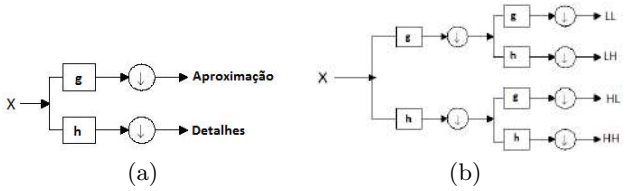

Figura 5: Processo de decomposição Wavelet (a) 1D e (b) 2D

Já a Transformada Wavelet descreve as propriedades de um sinal que muda durante o tempo [2]. Uma vantagem da Transformada Wavelet é que o tamanho da janela varia. Neste tipo de transformada, o sinal não é dividido em segmentos de tempo, mas em segmentos chamados de escala. Uma Transformada Wavelet permite analisar os dados em diferentes escalas, ou resoluções, as quais refletem as características mais ou menos detalhadas de um sinal. A análise de Wavelet consiste em adotar uma função-base e representar o sinal original como uma combinação linear de dilatações, contrações e translações (deslocamentos) sobre essa função-base. Existem várias famílias de funções-base Wavelets, dentre as quais destacam-se: Haar, Daubechies, Coiflets, Symlets e Morlet.

Ao invés de aplicarmos as *Funções Mães* para obtenção dos coeficientes de Wavelet, podemos aplicar um conjunto de pares de filtros FIR (resposta impulsiva finita) passa-alta H e passa-baixa G sobre o sinal. Após a aplicação destes filtros o sinal original é decomposto em duas partes, uma contendo os coeficientes de baixa frequência, representando os coeficientes de aproximação, e outra contendo os coeficientes de alta frequência, representado os coeficientes de detalhes.

De forma a utilizar a Transformada Wavelet em sinais 2D, tais como imagens, a ideia descrita acima pode ser expandida de forma a tratar as linhas e colunas da imagem como sinais independentes. Neste caso, o processo de decomposição é aplicado duas vezes seguidas, sendo que primeiro é realizado a análise nas linhas e a segunda nas colunas na imagem. Ao contrário da decomposição unidimensional, ao final desse processo geramos 4 sub-imagens diferentes contendo $\frac{M}{2} x \frac{N}{2}$ coeficientes, onde M é o número de linhas e N de colunas da imagem original. Essas sub-imagens são chamadas de LL, LH, HL e HH, onde LL representa a aproximação da imagem original e as outras contêm os detalhes nas direções horizontal, vertical e diagonal, respectivamente. As figuras 5(a) e 5(b) apresentam o processo de decomposição 1D e 2D.

Após a Transformada Wavelet ser aplicada em uma imagem, ou sinal, é possível realizar o procedimento novamente sobre os sinais resultantes, gerando assim um novo nível de decomposição. Uma das formas de realizar esse processo é através do esquema de pirâmide [18]. Nesse artigo, aplicamos Transformada Wavelet para a extração das características com 3 níveis de decomposição, sendo que os níveis 2 e 3 foram gerados através da decomposição da sub-imagem LL do nível anterior. As sub-imagens resultantes do nível 3 (LL_3, LH_3, HL_3 e HH_3) foram concatenadas e utilizadas como vetor de características.

5. MÁQUINAS DE VETORES SUPORTE

A Máquina de Vetores Suporte (SVM) é conhecida como uma das principais técnicas a ser utilizada em problemas de classificação. SVM já foi utilizada com sucesso em aplicações envolvendo sistemas biométricos baseado em íris [13] e

face [14]. Essa técnica é indicada para problemas em que os dados de entrada apresentam alta dimensionalidade [17], uma vez que a sua formulação não depende diretamente da dimensionalidade dos dados.

Dado um determinado problema de classificação com um conjunto de treinamento $\{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, com entrada $\mathbf{x}_i \in \Re^m$ e saída $y_i \in \{\pm 1\}$, considere um mapeamento não-Linear $\phi : \Re^m \longrightarrow \Re^n$, onde n é muito maior que m, ou seja, os dados de entrada são mapeados para um espaço com maior dimensionalidade [26]. Nesse novo espaço (chamado de espaço de característica), a SVM busca construir um hiperplano de separação ótima, $(\mathbf{w}^T\phi(\mathbf{x}) + b)$ onde a distância entre duas classes é maximizada. Neste caso, $\mathbf{w}$ pode ser definido como a combinação linear de $\phi(x)$, isto é, $\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \phi(\mathbf{x}_i)$ [4].

Na formulação original das SVMs proposta por Vapnik [28], o hiperplano ótimo $\mathbf{w}$ pode ser encontrado minimizando a seguinte função:

$$\min_{w,b,\xi_i} J(w, b, , \xi_i) = \frac{1}{2}(\mathbf{w}^T\mathbf{w}) + C\sum_{i=1}^{N} \xi_i, \qquad (1)$$

sujeita às restrições:

$$y_i[\mathbf{w}^T\phi(\mathbf{x_i}) + b] \geq 1 - \xi, i = 1, \ldots, N. \qquad (2)$$

onde a variável C representa um termo de compromisso entre a complexidade do modelo e o erro de treinamento, e $\xi_i, i = 1, \ldots, N$, são as variáveis de folga que medem a diferença entre a saída desejada e a saída da SVM. A solução para o problema de minimização da função-objetiva apresentada na equação 1, sujeito às restrições em 2, é dado pelo ponto de sela da função lagrangeana. O problema primal associado à minimização da função lagrangeana pode ser transformado em sua forma dual. O problema dual é dado por [28]:

$$\max_{\alpha} J(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i\alpha_j y_i y_j \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j) \qquad (3)$$

sujeito à $\sum_{i=1}^{N} \alpha_i y_i = 0$ e $0 \leq \alpha_i \leq C$, para $i = 1, \ldots, N$,onde $\alpha_i (i = 1, \cdots, N)$ são conhecidos como multiplicadores de Lagrange.

De forma a realizar o mapeamento $\phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$ na equação 3, não é necessário calcularmos $\phi(\mathbf{x}_i)$ ou $\phi(\mathbf{x}_j)$ de forma explícita, para isso podemos empregar as chamadas funções Kernel. Essas funções realizam o mapeamento não-Linear dos dados de entrada para um espaço de característica de maior dimensionalidade de forma eficiente [26], onde os dados são linearmente separáveis. Neste artigo utilizamos a função Kernel RBF, que pode ser descrita como:

$$K_{RBF}(\mathbf{x}_i, \mathbf{x}_j) = exp\left(-\frac{\| \mathbf{x}_i - \mathbf{x}_j \|^2}{2\sigma^2}\right). \qquad (4)$$

Após o processo de treinamento, as amostras que se encontram sobre a superfície de decisão apresentam valores de α_i, da equação 3, maiores que 0, sendo que essas amostras são chamadas de Vetores Suporte e são utilizadas para a definição da função de decisão do classificador [28]. Um determinado exemplo de teste $\mathbf{x} \in \Re^m$ é classificado ao final de acordo com a saída de:

$$f(n) = sign[\mathbf{w}^T\phi(\mathbf{x}) + b] = sign(\sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}, \mathbf{x_i}) + b)$$

.

6. EXPERIMENTOS COMPUTACIONAIS

A seguir, nós fornecemos detalhes a respeito do conjunto de dados usado nos experimentos e como estes foram configurados. Em seguida, nós apresentamos a taxa de reconhecimento obtida para os sistemas unimodais baseado em íris ou face e para os sistemas multimodais usando fusão no nível de característica.

6.1 Base de Dados SDUMLA-HMT

Há poucas bases de dados públicas na literatura. Neste artigo, nós utilizamos a base de dados *SDUMLA-HMT Multimodal Database*. Essa base de dados é pública e contém as seguintes modalidades biométricas: Face, Íris, Veias do Dedo, Caminhar e Impressão Digital de 106 pessoas diferentes. No nosso experimento foram utilizadas apenas as 50 primeiras pessoas. O objetivo desta redução foi reduzir o tempo de processamento gasto na construção dos sistemas biométricos unimodais e multimodais.

A base de dados de face contém 84 imagens coloridas por usuário. As imagens foram capturadas com o usuário em 12 poses diferentes e cada pose contém 7 diferentes orientações. No total, a base de dados contém 8904 imagens. Na Figura 6 podemos verificar alguns exemplos de imagens de face dessa base. Nos experimentos realizados foram selecionadas apenas 10 imagens contendo o usuário em posição frontal com a câmera, de forma a simular um ambiente controlado em um Sistema Biométrico baseado em Face, onde grandes variações em relação à pose e orientação não são esperadas.

Figura 6: Exemplos de imagens de face da Base de Dados SDUMLA-HMT.

A base de dados de Íris contém 10 imagens por usuário, sendo 5 imagens representando cada íris (esquerda e direita), e capturadas a uma distância de 6 - 32 cm do dispositivo utilizado. A imagem é armazenada em escala de cinza com dimensão de 768 × 576 pixels. A base de dados contém 1060 imagens ao todo. Na Figura 7 podem ser visto alguns exemplos das imagens de Íris dessa base.

Figura 7: Exemplos de imagens de íris da Base de Dados SDUMLA-HMT.

Para a combinação de Face e Íris foi selecionada, de forma aleatória, uma das imagens de face e uma das imagens de íris, em seguida foi aplicada uma das estratégias de combinação

descritas na Seção 3.1.

6.2 Pré-Processamento das Imagens

De forma a preparar as imagens para o processo de extração de características, alguns passos foram realizados previamente para fazer a remoção de possíveis informações irrelevantes das imagens a serem utilizadas.

Inicialmente, as imagens de face foram convertidas para escala de cinza. Em seguida, aplicamos o Algoritmo de Viola-Jones de forma a realizar a detecção da região da face na imagem de entrada e reduzir a interferência que o plano de fundo pode causar ao sistema. Após a detectada a região da face, a imagem resultante foi ajustada para a resolução de 128×128 pixels. Posteriormente, a Transformada Wavelet foi aplicada para a extração das características usando 3 níveis de decomposição.

Em relação às imagens de Íris, também foi aplicada uma técnica de detecção e segmentação. Neste caso, nós empregamos o Método de Wildes e Camus para detecção e o Modelo de Daugman para segmentação da íris. Neste caso, a imagem polar foi gerada com a mesma dimensão da imagem de face, isto é, 128×128 pixels. Em seguida a Transformada Wavelet também foi aplicada sobre essas imagens usando 3 níveis de decomposição.

Após a aplicação da Transformada Wavelet usando 3 níveis de decomposição, independente da característica biométrica, utilizamos as imagens LL_3, HL_3, LH_3 e HH_3 para compor o vetor de características que será utilizado pelo classificador. Na Tabela 1 apresentamos as funções Wavelets utilizadas e a quantidade de características geradas por cada função Wavelet.

Tabela 1: Funções Wavelets utilizadas e a quantidade de características extraída.

Função Wavelet	Quantidade de Características
Daubechies 2	1296
Daubechies 4	1936
Symlet 3	1600
Symlet 4	1936
Symlet 5	2116

6.3 Configuração dos Experimentos

Para a construção dos Sistemas Biométricos utilizamos as SVMs para compor o Módulo de Comparação. As SVMs são um classificador binário e não podem ser utilizada diretamente em problemas com múltiplas classes. De forma a superar essa limitação utilizamos a abordagem um-contra-um. Neste caso, $\frac{nc*(nc-1)}{2}$ SVMs diferentes foram treinados representando todas as possíveis combinações de classes, nc representa o número de classes. A saída é definida utilizando voto majoritário.

Além disso para garantir que as SVMs possam ser analisadas em relação ao seu melhor cenário é necessário que os seus hiperparâmetros sejam sintonizados adequadamente. Neste caso, o parâmetro C que controla a complexidade do classificador e o parâmetro relacionado com a função de kernel utilizada, neste caso o Kernel RBF (equação 4). O parâmetro σ do Kernel RBF controla a abertura do mapeamento não-linear que é realizado sobre os dados de entrada. Baseado em testes preliminares, o valor do parâmetro C foi fixado em 1000, enquanto que o valor do parâmetro σ foi definido como 2^i, com $i = 0, 1, \cdots, 9$.

Para melhor avaliar o desempenho do sistema biométrico construído foi utilizada validação cruzada 10-Fold. Nessa estratégia o conjunto de dados é separado em 10 partes diferentes, com a mesma quantidade de dados, onde uma parte é utilizada para teste e as 9 restantes são usadas no treinamento. Esse processo é repetido 10 vezes, a cada repetição uma parte diferente é escolhida como conjunto de teste. Este processo garante que todas as partes foram utilizadas para teste pelo menos uma única vez. O desempenho então foi avaliado de acordo com a taxa de acerto que o classificador obteve durante esse procedimento.

7. RESULTADOS

De forma a verificar a viabilidade e a capacidade dos Sistemas Biométricos Multimodais construídos usando as estratégias de fusão propostas na Seção 3.1, primeiramente realizamos a construção dos Sistemas Biométricos Unimodais usando Face ou Íris. Os resultados obtidos estão apresentados na Tabela 2. Para cada função Wavelet utilizada, apresentamos a taxa média de erro e o desvio padrão alcançado para o melhor cenário usando validação cruzada. Além disso, apresentamos o valor do parâmetro σ para o qual a SVM produziu melhor resultado.

Analisando os resultados alcançados utilizando somente os coeficientes de wavelet extraídos da Face, pode se perceber que, independente da função Wavelet utilizada, não foi possível obter uma taxa de erro menor que 10%. Neste caso, o melhor desempenho foi obtido usando a função Daubechies de ordem 2, que gerou uma taxa de erro de 16,20% com um valor de parâmetro do kernel (σ) igual a 128.

Com relação ao valor do parâmetro σ, podemos notar que, independente da função Wavelet utilizada, o melhor conjunto de valores de σ reside no intervalo de $[64, 128]$. Isto mostra que no caso da construção de um sistema biométrico envolvendo face, Transformada Wavelet e Kernel RBF, podemos escolher um valor dentro desse intervalo como um bom ponto de partida para o parâmetro do kernel. Comparando

Tabela 2: Resultados obtidos com os Sistemas Biométricos Unimodais usando Face e Íris.

Wavelet	**Face**		**Íris**	
	σ	**Taxa de Erro**	σ	**Taxa de Erro**
Daubechies 2	128	16,20%±6,69	256	7,80%±4,57
Daubechies 4	128	20,00%±5,81	64	10,00%±5,42
Symlet 3	128	17,40%±5,66	128	8,20%±3,46
Symlet 4	128	17,60%±5,94	64	9,80%±4,85
Symlet 5	64	19,20%±7,00	64	10,20%±6,43

as taxas de erro obtidas utilizando os coeficientes extraídos da íris com aquelas obtidas com relação à face, nota se que os sistemas baseados em íris alcançaram altas taxas de acerto. No melhor caso, foi obtido 7,80% de erro, com a função Daubechies de ordem 2. Para as funções Wavelet Symlet de ordem 3 e 4 a taxa de erro ficou abaixo de 10% de erro.

Já em relação ao valor do parâmetro σ podemos verificar que os melhores resultados para íris foram alcançados usando o parâmetro σ no intervalo $[64, 128]$, com exceção para o melhor cenário obtido que foi com o valor igual a 256 para função Daubechies de ordem 2. Apesar disso, podemos dizer que o intervalo $[64, 128]$ pode também ser utilizado como um ponto de partida para encontrar o melhor valor do parâmetro σ.

Os sistemas multimodais foram construídos utilizando as estratégias de fusão proposta na Seção 3.1, envolvendo a face

Tabela 3: Resultados obtidos com os Sistemas Biométricos Multimodais usando diferentes estratégias de fusão.

Função Wavelet	Estratégia 1		Estratégia 2		Estratégia 3		Estratégia 4	
	σ	Taxa de Erro	σ	Taxa de Erro	σ	Taxa de Erro	σ	Taxa de Erro
Daubechies 2	64	2,00%±1,63	128	1,60%±1,26	128	5,40%±2,31	64	1,60%±1,57
Daubechies 4	64	2,00%±2,66	128	3,20%±2,34	64	5,20%±3,15	128	2,60%±2,67
Symlet 3	64	2,60%±2,11	64	2,60%±2,83	128	4,20%±2,74	64	2,60%±2,11
Symlet 4	64	2,80%±2,69	64	3,00%±1,69	128	3,60%±2,63	64	1,60%±1,83
Symlet 5	64	2,40%±1,57	64	2,20%±1,47	128	4,00%±3,12	128	3,00%±2,35

e íris. Os resultados obtidos para cada estratégia são apresentados na Tabela 3. A primeira estratégia de fusão utiliza os coeficientes LL_3 extraídos da sub-imagem de face e os coeficientes LH_3, HL_3 e HH_3 extraídos da sub-imagem de íris concatenados para gerar o vetor de característica. Analisando a Tabela 3, pode se notar que o desempenho obtido com essa abordagem conseguiu superar os melhores resultados obtidos pelos sistemas unimodais. Para esta estratégia, nós obtemos uma taxa média de erro menor que 3%, independente da função Wavelet utilizada.

O melhor resultado para a primeira estratégia foi alcançado usando a função Daubechies de ordem 2 e 4, na qual obtivemos uma taxa de erro de 2%. Para esta taxa de erro, o menor desvio padrão foi alcançado usando a função Daubechies de ordem 4. Usando a primeira estratégia, foi possível superar o desempenho alcançado pelos sistemas unimodais baseado em íris. Em relação à variação do parâmetro σ pode-se observar que em todos os experimentos para essa estratégia, os melhores resultados foram alcançados usando o valor igual a 64. Este valor ficou dentro do mesmo intervalo obtido para os sistemas biométricos unimodais.

A segunda estratégia utiliza os coeficientes LL_3 extraídos da sub-imagem da íris concatenada com os coeficientes das sub-imagens LH_3, HL_3 e HH_3 da face dos usuários. Novamente foi possível superar o desempenho alcançado pelos sistemas biométricos unimodais.

Usando a função Daubechies de ordem 2, a segunda estratégia produziu resultados superiores aos alcançados pela primeira e terceira estratégia. Sendo que, para o melhor cenário, foi possível obter uma taxa de erro média igual a 1,60%. Em relação ao parâmetro σ é possível verificar que o mesmo intervalo obtido pelos sistemas unimodais ($[64, 128]$) foi alcançado usando a segunda estratégia. Sendo que o melhor resultado foi obtido usando o valor de σ igual a 128.

A terceira estratégia utiliza a média dos coeficientes das sub-imagens LL_3 da íris e da face, e o máximo dos coeficientes das sub-imagens LH_3, HL_3 e HH_3 para compor o vetor de características. Semelhante às outras estratégias de fusão, foi possível superar o desempenho dos sistemas unimodais, independente da função Wavelet adotada.

Ao contrário da outras estratégias utilizadas e dos sistemas biométricos unimodais, essa estratégia obteve o melhor desempenho usando a função Wavelet Symlet de ordem 4. Neste caso, foi possível alcançar uma taxa de erro igual a 3,60% com o valor de σ igual a 128. Em relação ao valor do parâmetro σ, o mesmo intervalo adotado por todas as estratégias e os sistemas unimodias foi também obtido por esta estratégia, isto é $[64, 128]$. Por outro lado, o desempenho obtido com essa técnica foi o pior entre todas as estratégias de fusão adotada.

A última estratégia de fusão investigada aqui utiliza o mínimo dos coeficientes das sub-imagens LL_3 da íris e da face concatenados com o máximo dos coeficientes das sub-imagens LH_3, HL_3 e HH_3 da íris e face. Essa estratégia forneceu o melhor resultado dentre todas as estratégias de fusão adotadas, com exceção da segunda estratégia. Entretanto, para algumas funções wavelets, o desempenho foi pior quando comparado as outras estratégias, conforme pode ser observado na Tabela 3. O melhor cenário foi obtido usando a função Daubechies de ordem 2,neste caso foi alcançado uma taxa de erro igual a 1,60%. Esse resultado também foi obtido usando a função Symlet de ordem 4, porém com um desvio padrão maior.

Analisando os valores do parâmetro σ, pode se verificar que os melhores resultados em termos de taxa de erro de reconhecimento foram alcançados com valores de σ no intervalo $[64, 128]$. Esta mesma faixa de valores foi obtida para as outras estratégias de fusão e para os sistemas unimodais. Isto demonstra que esta faixa de valores contém bons candidatos para os valores de σ. Logo, ao utilizarmos a transformada Wavelet com 3 níveis de decomposição e imagens de entrada de dimensionalidade de 128×128, um bom valor de σ pode ser escolhido neste intervalo.

De forma geral, podemos dizer que os sistemas biométricos multimodais, independente da estratégia de fusão, podem alcançar taxas de erro menores que aquelas obtidas pelos sistemas unimodais. No melhor cenário, os sistemas multimodais alcançaram taxa de erro igual a $1,60\%$, enquanto os sistemas unimodais alcançaram taxa de erro igual a $7,80\%$. Com relação a função Wavelet, independente do tipo de sistema (unimodal ou multimodal) e da estratégia de fusão, os melhores resultados foram alcançados com função Wavelet Daubechies de ordem 2. Por outro lado, a estratégia de fusão parece ser uma decisão importante na construção de sistemas multimodais. No caso das estratégias de fusão investigadas a taxa de reconhecimento incorreta variou de 1.60% a 2.0%. Já com relação ao parâmetro do kernel, apesar deste ser importante tanto para os sistemas unimodais quanto para os multimodais, para todas as estratégias o melhor resultado foi alcançado para valores de *sigma* no intervalo $[64, 128]$. Acreditamos que este intervalo seja um bom ponto de partida para escolha do valor deste parâmetro.

8. CONCLUSÃO

Este trabalho realizou um estudo comparativo, em termos de taxa de reconhecimento incorreto, entre os sistemas biométricos unimodais e multimodais. As modalidades biométricas empregadas foram face e íris. Para a construção dos sistemas biométricos multimodais foi utilizada a fusão no nível de característica usando quatro estratégias diferentes. Para extração de característica foi utilizada a Transformada Wavelet.

Independente do tipo de estratégia de fusão, função Wavelet e modalidade biométrica, os sistemas biométricos multimodais alcançaram desempenho superior aos obtidos pelos sistemas unimodais. A estratégia de fusão parece ser um

ponto importante a ser levantado na construção de sistemas multimodais.

Como próximos passos, serão investigadas outras técnicas de extração de característica e qual o impacto dessas quando empregadas em um sistema multimodal. Além disso, serão investigadas outras formas de fusão, como por exemplo, fusão no nível sensorial e nível de decisão, e outros tipos de classificadores.

9. REFERÊNCIAS

[1] A. Batool and A. Tariq. Computerized system for fingerprint identification for biometric security. In *Multitopic Conference (INMIC), 2011 IEEE 14th International*, pages 102–106. IEEE, 2011.

[2] C. S. C. S. Burrus, R. A. Gopinath, and H. Guo. *Introduction to wavelets and wavelet transforms : a primer*. Upper Saddle River, N.J. Prentice Hall, 1998.

[3] T. Camus and R. Wildes. Reliable and fast eye finding in close-up images. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 389–394 vol.1, 2002.

[4] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 1 edition, 2000.

[5] J. Daugman. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14:21–30, 2002.

[6] K. Dharavath, F. Talukdar, and R. Laskar. Study on biometric authentication systems, challenges and future trends: A review. In *Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on*, pages 1–7, Dec 2013.

[7] V. Divyaloshini and M. Saraswathi. Performance evaluation of image fusion techniques and its implementation in biometric recognition. *International Journal of Technology Enhancements and Emerging Engineering Research*, 2(3), 2014.

[8] L. Flom and A. Saffir. Iris recognition system. U.S Patent 4 641 394, 1987.

[9] L. Hong and A. Jain. Integrating faces and fingerprints for personal identification. *IEEE transactions on pattern analysis and machine intelligence*, 20:1295–1307, 1998.

[10] A. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[11] A. K. Jain, P. Flynn, and A. A. Ross. *Handbook of Biometrics*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[12] V. Kabeer, T. M. Thasleema, and N. K. Narayanan. Face recognition using state space parameters and k-nn classifier. In *Innovations in Information Technology, 2007. IIT '07. 4th International Conference on*, pages 476–480, Nov 2007.

[13] H. F. Liau and D. Isa. Feature selection for support vector machine-based face-iris multimodal biometric system. *Expert Systems with Applications*, 38(9):11105 – 11111, 2011.

[14] Z. Liu, L. Zhang, and L. Zhu. An improved face recognition method based on gabor wavelet transform and svm. In *Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium on*, volume 1, pages 378–381, Oct 2012.

[15] M. A. Lone, S. Zakariya, and R. Ali. Automatic face recognition system by combining four individual algorithms. In *Computational Intelligence and Communication Networks (CICN), 2011 International Conference on*, pages 222–226. IEEE, 2011.

[16] A. Lourenço, H. Silva, and A. Fred. Ecg-based biometrics: A real time classification approach. In *Machine Learning for Signal Processing (MLSP), 2012 IEEE International Workshop on*, pages 1–6. IEEE, 2012.

[17] A. Lourenço, H. Silva, and A. Fred. Ecg-based biometrics: A real time classification approach. In *Proceedings of the 22nd IEEE International Workshop on Machine Learning for Signal Processing*, 2012.

[18] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(7):710–732, 1992.

[19] V. M. Mane and D. V. Jadhav. Review of multimodal biometrics: Applications, challenges and research areas. *International Journal of Biometrics and Bioinformatics (IJBB)*, 2009.

[20] M. Mazloom and S. Ayat. Combinational method for face recognition: Wavelet, pca and ann. In *Digital Image Computing: Techniques and Applications (DICTA), 2008*, pages 90–95, Dec 2008.

[21] M. Mazloom, S. Kasaei, and H. A. Neissi. Construction and application of svm model and wavelet-pca for face recognition. In *Proceedings of the 2009 Second International Conference on Computer and Electrical Engineering - Volume 01*, ICCEE '09, pages 391–398, Washington, DC, USA, 2009. IEEE Computer Society.

[22] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles'. The electroencephalogram as a biometric. In *Electrical and Computer Engineering, 2001. Canadian Conference on*, volume 2, pages 1363–1366 vol.2, 2001.

[23] H. Proenca and L. Alexandre. Iris recognition: An analysis of the aliasing problem in the iris normalization stage. In *Computational Intelligence and Security, 2006 International Conference on*, volume 2, pages 1771–1774, 2006.

[24] D. M. Rankin, B. W. Scotney, P. J. Morrow, and B. K. Pierscionek. Iris recognition failure over time: The effects of texture. *Pattern Recogn.*, 45(1):145–150, Jan. 2012.

[25] A. Ross and A. K. Jain. Multimodal Biometrics: an overview. pages 1221–1224, 2004.

[26] B. Schölkopf and A. J. Smola. *Learning with kernels : support vector machines, regularization, optimization, and beyond*. Adaptive computation and machine learning. MIT Press, 2002.

[27] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.

[28] V. N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.

[29] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Comp. Vision*, 57(2):137–154, 2004.