

Reconhecimento Automático de Gestos da Língua Brasileira de Sinais utilizando Visão Computacional

Kleber Padovani de Souza
Universidade Católica Dom Bosco
Grupo de Pesquisa em Engenharia e Computação
Campo Grande - MS, Brasil
kleber.padovani@uol.com.br

Jéssica Barbosa Dias
djessy@gmail.com

Hemerson Pistori
pistori@ucdb.br

Resumo

Hidden Markov Models (HMMs) have been frequently used in speech recognition, what suggested their possible use in gesture recognition. Traditional human-machine interfaces are not suitable for people with certain disabilities, specially auditory and motor disabilities, and the use of computers to automatically recognize gestures can bring important benefits to this community. This work presents a HMM-based system integrated in a platform to support the development of visually guided applications capable of recognizing human gestures in sequences of images achieving an accuracy rate of 99% for training samples and 80% for testing samples in a subset of Brazilian Sign Language gestures.

1. Introdução

O reconhecimento automático de gestos humanos pode ser utilizado em várias aplicações úteis à sociedade, como em ambientes virtuais, em sistemas de segurança e supervisão, em ferramentas de suporte médico, no suporte à interação homem-máquina através de movimentos e em muitas outras. Existem duas abordagens que são comumente utilizadas para realizar este reconhecimento.

A primeira é através do uso de alguns dispositivos de apoio, que auxiliam na coleta de informações relevantes ao processo de classificação, como as luvas de dados e *body suits*. Esta abordagem, apesar de simplificar e, geralmente, aumentar o nível de precisão na extração de características das imagens, traz consigo algumas desvantagens, como a limitação dos movimentos do sinalizador¹ e o desconforto, causado pela necessidade de se vestir o equipamento. Em

muitos casos, existe rejeição por parte dos usuários a este tipo de recurso, devido a esta característica intrusiva [9].

A segunda abordagem é baseada em visão computacional, em que a classificação dos gestos é alcançada obtendo-se apenas imagens como fonte de informação, eliminando o incômodo implícito na abordagem anterior. No entanto, o custo computacional de sistemas baseados em visão, geralmente é maior, devido à necessidade de processamento das imagens. Todavia, no intuito de facilitar a implementação de sistemas com interfaces guiadas por sinais visuais, foi criada uma plataforma de códigos-fonte abertos, denominada SIGUS [14], que minimiza a complexidade desta tarefa.

Neste trabalho, foram investidos esforços no desenvolvimento de um sistema de reconhecimento automático de gestos da Língua Brasileira de Sinais, baseado na plataforma SIGUS, utilizando modelos de Markov ocultos (ou *Hidden Markov Models* - HMMs) no estágio de classificação. Os HMMs são estruturas amplamente utilizadas na modelagem de problemas com variações temporais e estão entre os métodos mais utilizados nas pesquisas da área de reconhecimento de gestos [4, 12]².

A descrição deste trabalho foi organizada em 4 partes. Na seção seguinte é feita uma breve revisão sobre a técnica central do trabalho, os modelos de Markov ocultos (HMM). A seção 3 explica detalhadamente a forma em que o conceito foi aplicado ao reconhecimento de gestos e cita alguns detalhes sobre as implementações. Por fim, as seções 4 e 5 apresentam os experimentos e resultados obtidos, faz uma análise breve e geral sobre o tema discutido neste trabalho e sugere possíveis trabalhos futuros.

¹ Os termos sinalizador e colaborador são utilizados para referenciar aos usuários que executaram os gestos para obtenção das amostras (vídeos) dos gestos.

² Apesar de algumas literaturas considerarem posturas (ou configurações) como gestos estáticos, neste trabalho, o termo gesto se refere apenas a gestos dinâmicos.

2. Modelos de Markov Ocultos

Nas últimas décadas, diversos pesquisadores da área de reconhecimento de fala têm utilizado com sucesso os modelos de Markov ocultos como mecanismos de classificação em seus trabalhos [16, 20, 8, 1]. Recentemente, eles têm sido empregados em aplicações de visão computacional que envolvem reconhecimento de padrões, como no reconhecimento de letras manuscritas [10, 19, 22] e na classificação de gestos em seqüências de imagens [12, 11, 5, 21].

Um modelo de Markov oculto é similar a uma Cadeia de Markov[7], porém, os elementos dos conjuntos de observações não pertencem ao conjunto de estados, mas a outro conjunto, chamado conjunto de símbolos, tornando a seqüência de estados percorrida oculta ao observador. Esses símbolos são gerados pelos estados e, em cada instante de tempo, é gerado um símbolo pelo conjunto de estados, no entanto, o estado gerador não é conhecido. Portanto, é acrescentada uma camada estocástica na cadeia de Markov para se obter um modelo de Markov oculto.

A primeira camada estocástica é uma cadeia de Markov de primeira ordem, porém, não é diretamente observável como nas cadeias de Markov, em que cada estado é uma possível observação³. A segunda camada estocástica é um conjunto de probabilidades que indica, para cada estado, as probabilidades de emissão de cada símbolo do modelo.

Como dito, a seqüência de estados percorrida em um modelo, dada uma seqüência de observações, é oculta ao observador. Ou seja, tendo uma seqüência de observações (conjunto ordenado de símbolos), não se tem acesso à seqüência de estados percorrida pelo modelo para geração desta seqüência, mas somente uma função probabilística deste caminho, e, por isso, o modelo é chamado de modelo de Markov oculto. A seguir, serão descritos alguns termos adotados que são comumente encontrados em textos da área e ao longo deste trabalho.

1. Um HMM é representado pelo símbolo λ ;
2. Os estados do modelo são denotados pelo conjunto $S = \{s_1, s_2, \dots, s_N\}$, de tamanho N ;
3. Os símbolos reconhecidos pelo modelo estão contidos no conjunto $V = \{v_1, v_2, \dots, v_M\}$, de tamanho M , também conhecido como alfabeto do modelo;
4. Uma seqüência de observações é denotada pelo conjunto ordenado $O = \{o_1, o_2, \dots, o_T\}$, composto de T elementos quaisquer do conjunto V , em que T e o_t representam, respectivamente, o tamanho da seqüência e

³ O termo “observações” recebe diferentes significados quando aplicados a processos de Markov e a modelos de Markov ocultos. No primeiro, essas observações são os estados percorridos pelo modelo, enquanto que nos modelos de Markov ocultos são os símbolos gerados nos estados do modelo, porém, a seqüência de estados não é conhecida.

o símbolo observado no instante t da seqüência, tal que $1 \leq t \leq T$;

5. Quando conhecida, uma seqüência de estados para determinada observação é representada pelo conjunto $Q = \{q_1, q_2, \dots, q_T\}$, composto por T elementos de S , em que q_t representa o estado no instante t da seqüência de observações de tamanho T ;
6. O vetor $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$, com um valor probabilístico para cada um dos N estados do modelo, sendo π_{estado} a probabilidade de *estado* ser o gerador do primeiro símbolo de qualquer seqüência de observações gerada pelo modelo;
7. A matriz $A_{N \times N}$, cujos elementos são referenciados na forma $a_{(origem, destino)}$ e representam a probabilidade de transição do estado *origem* para o estado *destino*;
8. A matriz $B_{N \times M}$, cujos elementos são representados como $b_{estado(símbolo)}$ e representam a probabilidade de um *estado* gerar determinado *símbolo*;
9. A probabilidade da seqüência de observação O ter sido gerada pelo modelo λ é representada por $P(O|\lambda)$.

Os elementos de π , A e B podem ser os índices dos elementos dentro dos conjuntos ou seus próprios nomes. Por exemplo, a probabilidade do estado s_i ser inicial pode ser descrita como π_i ou π_{s_i} e a probabilidade do estado s_i ser inicial pode ser descrita como π_i ou π_{s_i} . Da mesma forma, $b_i(v_j)$ indica a probabilidade do estado s_i gerar o símbolo v_j e $b_{s_x}(y)$ indica a probabilidade do estado s_x gerar o símbolo v_y .

Para melhor compreensão, a Figura 1 ilustra um exemplo de HMM com três estados (1, 2 e 3) e dois símbolos (a e b). A matriz A seria composta pelas probabilidades de transição entre estados (e.g. $a_{(1,2)} = 0.4$, $a_{(3,2)} = 0.6$, etc.) e a matriz B pelas probabilidades de emissão de símbolos de cada estado (e.g. $b_3(a) = 0.7$, $b_2(b) = 0.9$, etc.), restando apenas o vetor π , que conteria as probabilidades iniciais dos estados 1, 2 e 3. As possíveis seqüências de observações para este modelo seriam todas compostas pelos símbolos a e b (e.g. $O = \{a, a, b, a, b\}$) de qualquer tamanho T , tal que $T \geq 1$.

2.1. Incógnitas principais

Existem três incógnitas implícitas nos HMMs, cujas soluções contribuem para o funcionamento eficaz das aplicações que o utilizam [13], sendo elas a avaliação da observação, a melhor seqüência de estados da observação e o treinamento dos modelos⁴. As

⁴ As três incógnitas descritas são citadas por alguns autores como os três “problemas básicos” dos HMMs, como o problema da avaliação, o problema do melhor caminho e o problema do treinamento.

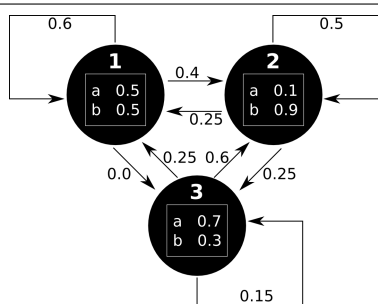


Figura 1. Exemplo de HMM.

três incógnitas e suas soluções serão brevemente discutidas. Para maiores detalhes, consulte [16].

2.1.1. Avaliação da observação O primeiro problema se refere à descoberta da probabilidade de uma determinada seqüência de observações O ter sido gerada por um específico modelo λ . Este tipo de situação pode ser muito freqüente nas aplicações de HMM. Como exemplo, no reconhecimento de voz, ao se produzir um fonema qualquer, esta entrada pode ser classificada como pertencente ao modelo que indicar a maior probabilidade ($P[O|\lambda]$); outro exemplo seria no reconhecimento de comportamento de animais em seqüências de imagens, em que o modelo que apresentar a maior probabilidade indica a reação do animal perante uma situação ou um conjunto de estímulos.

A maneira mais “simples” de calcular a probabilidade de determinada seqüência de observação ter sido gerada por um modelo é através da verificação de todas as seqüências de estados de tamanho T (número total de observações) possíveis, e, posteriormente, calcular suas probabilidades. No entanto, este cálculo pode ser computacionalmente impraticável devido à quantidade de operações matemáticas necessárias. A fim de minimizar o número de operações, podem ser utilizados dois métodos, chamados *forward* e *backward* [16]. Ambos utilizam estratégia de programação dinâmica para minimizar o custo computacional do cálculo. Contudo, vale lembrar que os procedimentos *forward* e *backward* são independentes e apenas um deles é necessário para encontrar a probabilidade desejada.

2.1.2. Melhor seqüência de estados da observação A segunda incógnita se refere à busca da melhor seqüência de estados percorrida pelo modelo para uma determinada seqüência de observações. Ou seja, a seqüência de estados que produz a maior probabilidade final para uma determinada seqüência de observações dentre todas as seqüências de estados possíveis no modelo. Um exemplo de situação em que esta incógnita ocorre é na transformação de fala em texto, em que os sinais acústicos são as observações do sistema e as letras são os estados do modelo. Assim, dado um sinal acústico, procura-se a seqüência de estados

mais provável, sendo esta a seqüência de letras equivalente ao sinal. Geralmente, esta incógnita é resolvida por meio do algoritmo de Viterbi, que também utiliza a técnica de programação dinâmica.

2.1.3. Treinamento do modelo A terceira incógnita é a realização do ajuste dos parâmetros (A, B, π) do modelo a partir de uma ou mais seqüências de observações. A seqüência O utilizada para este ajuste é chamada de seqüência de treinamento, pois é utilizada para treinar o HMM. Não existe uma maneira conhecida de realizar este ajuste para resolver analiticamente o modelo que maximize a probabilidade da(s) seqüência(s) de observações, porém, é possível escolher o modelo que sua probabilidade seja localmente maximizada usando um procedimento iterativo.

A existência de um processo que ajusta iterativamente e automaticamente os parâmetros do modelo com base nas seqüências de observações é a principal força dos HMMs [6]. Este processo é executado pelo algoritmo Baum-Welch [16], que se trata de uma especialização do algoritmo EM - *Expectation-Maximization* [2, 15] aplicada aos HMMs. Um exemplo de aplicação desta incógnita ocorre nos treinamentos dos sistemas que utilizam HMMs no processo de classificação, em que, dadas as amostras das classes que se deseja reconhecer, o algoritmo ajusta os modelos, a fim de se otimizar o reconhecimento de outras seqüências de observações similares.

3. Desenvolvimento

Com auxílio do dicionário trilingüe [17, 18], foram selecionados alguns gestos que são utilizados na formação das sentenças da Língua Brasileira de Sinais, LIBRAS, e construído um banco de amostras (vídeos) de gestos [3]. Para cada gesto foi desenvolvido um respectivo modelo de Markov oculto.

Para a construção desses modelos foram escolhidas as posturas mais marcantes dos gestos, sendo que cada uma destas posturas se relaciona diretamente a um estado do modelo. Uma postura é uma configuração estática, sem movimento, enquanto o gesto é dinâmico, ou seja, possui movimento. Por exemplo, a foto de uma mão e a filmagem de uma cabeça se deslocando da esquerda para a direita são exemplos de postura e gesto, respectivamente.

Primeiramente, foram definidas as características que, juntas, auxiliam a discriminar as posturas e, conseqüentemente, os gestos executados pelos sinalizadores. As características foram escolhidas com base nas descrições e padronizações contidas em [17, 18] e na análise dos gestos escolhidos, observando quais combinações discriminavam visualmente as posturas e os gestos selecionados de maneira única e, posteriormente, foram associadas as pos-

turas marcantes aos estados e destacadas as transições entre essas posturas.

Para obtenção das probabilidades de algumas componentes dos HMMs foi realizada uma análise visual, em que um especialista analisou cada quadro (postura) dos vídeos (gestos) e a relacionou a um estado do respectivo modelo, obtendo as seqüências de estados para cada amostra de gesto e, conseqüentemente, as probabilidades das matrizes de ocorrência inicial (π) e de transição de estados (A). Durante a análise, como apenas as posturas denominadas como marcantes têm um respectivo estado no modelo, algumas posturas são “ignoradas”⁵, sendo substituídas nessa correlação pela postura marcante mais próxima.

Por exemplo, a Figura 2 ilustra as posturas marcantes para execução do gesto referente a “derrame-cerebral”, em que a mão direita deve estar em M⁶, com a palma virada para a esquerda e com os dedos apontando para cima, tocando o lado direito da testa. Em seguida, cruzar os braços em frente à cabeça, com as duas mãos fechadas e com as palmas para dentro, e mover as mãos para os lados opostos com força.

Como foram consideradas três posturas como marcantes, conseqüentemente, o modelo construído para a representação deste gesto conterà três estados. Desta forma, as posturas que são diferentes dessas, como as posturas necessárias para transitar de uma postura marcante para outra, são aproximadas para uma das posturas marcantes durante a análise.



Figura 2. Posturas denominadas marcantes para o gesto “derrame cerebral”.

A matriz de emissão de símbolos (B) se relaciona com o sinal real obtido, que, no caso, são as informações obtidas através da seqüência de imagens dos gestos. Para obtenção dessas características, as imagens eram submetidas ao processo de segmentação e, posteriormente, à extração de características. Para detecção de pele nas imagens, foram utilizadas árvores de decisão e, em seguida, foi realizado pre-

5 As posturas contidas no gesto que são classificadas como desconhecidas para o modelo são aproximadas pelo sistema a uma das posturas marcantes, obedecendo a determinado critério de classificação.

6 A descrição detalhada de todos os gestos selecionados podem ser obtidas no dicionário trilingüe da Língua de Sinais Brasileira, de Fernando Capovilla.

enchimento de orifícios e exclusão de regiões irrelevantes. A partir das imagens segmentadas, foram extraídas características da face e das duas mãos do sinalizador, como área, centro de massa, excentricidade, etc.

Como a estrutura dos HMMs utilizados era discreta, ou seja, as observações compatíveis com o modelo estão em domínio discreto, os atributos oriundos do processo de extração de características eram submetidos a uma etapa de quantização e a combinação dos valores discretos dos atributos produzia um símbolo do modelo e, por conseqüência, a seqüência de observações da amostra. Assim, tendo as seqüências de estados e observações, foram obtidas as probabilidades de emissão de símbolos para cada estado dos modelos e, desta forma, os modelos iniciais referentes aos gestos estavam completos, possuindo suas triplas (π , A , B) necessárias.

Para reconhecer as amostras dos gestos, elas são enviadas a um gerenciador de modelos, que considera as probabilidades de geração de cada modelo para determinada seqüência de observações. Alternativamente, este gerenciador possui a opção de realizar a contagem de símbolos reconhecidos pelos modelos, ou seja, os símbolos da seqüência de observações que têm probabilidade diferente de zero em pelo menos um estado do modelo. Desta maneira, seqüências com posturas que, por curtos intervalos de tempo não seguem um comportamento esperado pelo gesto, podem ser reconhecidas de maneira apropriada.

Considerando que o sistema possui um conjunto de símbolos muito extenso (cerca de 210 bilhões), existem diversas seqüências de observações possíveis para um mesmo gesto. Dessa forma, o algoritmo distribui em cada modelo, de maneira equiprovável, uma contribuição irrisória aos símbolos que não estavam contidos nas amostras utilizadas para obtenção dos modelos. Do mesmo modo, o gerenciador distribui probabilidades mínimas às condições inexistentes às transições e ocorrências iniciais dos estados, que, por serem muito pequenas, não interferem consideravelmente na classificação e podem auxiliar na correção de erros ou restrições provenientes da fase de treinamento.

Do mesmo modo, o gerenciador de modelos implementado executa a mesma distribuição equiprovável a cada iteração do algoritmo Baum-Welch, com intuito de possibilitar o aumento de probabilidade dos símbolos não reconhecidos por cada modelo. Porém, os valores dessa distribuição e a ativação de sua execução é parametrizada, possibilitando o procedimento normal, sem adaptações. Entre os valores parametrizados, está o índice de estabilidade entre iterações, que determina o momento de encerramento das iterações de reestimação.

Outra característica peculiar do gerenciador é a alternativa que habilita a mesclagem de amostras de treinamento,

que possibilita a execução do método reestimador em várias seqüências de observação em cada iteração, considerando todas as observações antes de maximizar as probabilidades dos modelos.

Como as probabilidades são muito pequenas devido ao número de símbolos, algumas não seriam suportadas em variáveis comuns da linguagem de programação escolhida, por isso, foi implementada uma estrutura especial, que apresenta suporte ao armazenamento de números dessa natureza.

Os códigos implementados para realização de todas as etapas do sistema reconhecedor são abertos e foram desenvolvidos na linguagem Java, utilizando bibliotecas de processamento de imagens (ImageJ) e mineração de dados (Weka) e estão em processo de incorporação na plataforma SIGUS para utilização em outras aplicações.

4. Experimentos e Resultados

Para a realização das análises e experimentos no sistema, foi criado um banco de imagens referentes aos gestos selecionados [3]. A captura das imagens que compõem este banco foi realizada em um ambiente com fundo estático e uniforme, ou seja, não existiam objetos atrás do sinalizador e o único objeto com movimento era ele próprio. Para isto, foi utilizada uma câmera digital⁷ para obtenção das seqüências de imagens, sendo que esta se manteve fixada em frente ao sinalizador e imóvel durante as gravações.

A coleta foi organizada em ciclos para evitar que o sinalizador executasse o mesmo gesto mais de uma vez, com o intuito de aumentar a naturalidade na execução dos gestos, explorando variações possíveis. Foram capturadas imagens de 3 colaboradores para os experimentos, em que cada um executou sete ciclos de 47 gestos. Com isso, o banco de imagens dedicado à experimentação possuía 21 amostras para cada gesto, totalizando 987 arquivos de vídeo, em que não foi utilizado recurso algum de apoio para o rastreamento das mãos e da face, como luvas de dados ou câmeras de infra-vermelho.

Para realização dos experimentos, foram divididas as amostras em duas classes, treinamento e teste, e os experimentos em 2 grupos, sendo que as amostras de treinamento são as mesmas utilizadas para obtenção dos modelos iniciais. Dentro de cada classe são realizados dois experimentos: com HMMs constituídos das probabilidades alcançadas por fórmula fechada utilizando os dados obtidos durante a análise visual e com HMMs submetidos ao processo de reestimação, utilizando as mesmas amostras de treinamento.

A primeira classe experimenta a classificação das amostras de treinamento, que são as mesmas utilizadas para a

obtenção dos modelos iniciais e a segunda as amostras restantes. A base de treinamento possui um conjunto total de 987 amostras de gestos, sendo que 564 dessas amostras são destinadas ao treinamento (e testes) e as restantes apenas para testes. Em ambas as abordagens, com modelos iniciais e reestimados, o sistema obteve taxa de acerto de 99.11% no primeiro grupo e de 80.1% nas amostras de teste, que são as amostras que não foram utilizadas para obtenção dos modelos iniciais.

5. Considerações Finais

Neste trabalho investiu-se na utilização de uma técnica bastante utilizada no reconhecimento de fala e muito bem conceituada, porém, aplicada ao reconhecimento de gestos: os modelos de Markov ocultos (*HMM*). O domínio deste sistema reconhecedor é um conjunto de 47 gestos, extraídos do dicionário Trilingüe da Língua Brasileira de Sinais [17, 18]. Estes gestos foram analisados e capturados através de câmeras digitais e armazenados em um banco de imagens. Foram analisadas as variáveis visuais que melhor caracterizavam os gestos, ou seja, um conjunto de atributos que, em conjunto, descrevem cada gesto de maneira distinta. De fato, estas características não descrevem especificamente o gesto, mas sim, cada postura do gesto.

Os resultados obtidos foram satisfatórios, no entanto, podem ser analisadas alternativas que possam maximizar os percentuais de acerto. Como exemplo, a análise detalhada da invariância na taxa de acerto antes e após a reestimação dos modelos talvez possa auxiliar na descoberta de características relevantes ao processo de classificação.

Uma das melhorias que poderia ser adicionada ao trabalho, seria a incorporação de experimentos em imagens capturadas em tempo real, pois os experimentos realizados no sistema se limitaram ao domínio de imagens armazenadas em disco. Além disso, outra abordagem que enriqueceria o projeto, seriam experimentos com colabores que não estão presentes na parcela de treinamento, para constatar se o sistema suporta independência de usuários.

Outra abordagem interessante seria uma análise comparativa entre o desempenho de classificação do sistema aplicando esta mesma metodologia para a construção dos modelos em relação a modelos gerados aleatoriamente e reestimados através do algoritmo Baum-Welch. Essa análise demonstraria a relevância da identificação das seqüências de estados nas amostras, como foi realizado neste trabalho durante a análise visual.

6. Agradecimentos

Este trabalho recebeu apoio financeiro da Universidade Católica Dom Bosco, UCDB, da Agência Financiadora de

⁷ Câmera digital, marca Sony, modelo DSC-F828.

Estudos e Projetos, FINEP e da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, FUNDECT. O primeiro autor deste trabalho foi bolsista ITI-CNPq e um dos co-autores é bolsista de Produtividade em Desenvolvimento Tecnológico e Extensão Inovadora do CNPq.

Referências

- [1] T. Aarnio. Speech recognition with hidden markov models in visual communication. Master's thesis, University of Turku, 1999.
- [2] F. Dellaert. The expectation maximization algorithm. Technical Report GIT-GVU-02-20, GVU, February 2002.
- [3] J. B. Dias, K. P. Souza, and H. Pistori. Conjunto de treinamento para algoritmos de reconhecimento de libras. *II Workshop de Visão Computacional, São Carlos*, October 2006.
- [4] S. Eickeler, A. Kosmala, and G. Rigoll. Hidden Markov Model Based Continuous Online Gesture Recognition. In *Int. Conference on Pattern Recognition (ICPR)*, pages 1206–1208, Brisbane, 1998.
- [5] S. Eickeler, A. Kosmala, and G. Rigoll. Hidden Markov Model Based Continuous Online Gesture Recognition. In *Int. Conference on Pattern Recognition (ICPR)*, pages 1206–1208, Brisbane, 1998.
- [6] A. El-Yacoubi, R. Sabourin, M. Gilloux, and C. Y. Suen. *Off-line handwritten word recognition using hidden Markov models*, pages 191–230. CRC Press, Inc., Boca Raton, FL, USA, 1999.
- [7] C. M. Grinstead and L. J. Snell. *Introduction to Probability*. American Mathematical Society, July 1997.
- [8] X. D. Huang, H. W. Hon, and K. F. Lee. Large-vocabulary speaker-independent continuous speech recognition with semi-continuous hidden markov models. In *HLT '89: Proceedings of the workshop on Speech and Natural Language*, pages 276–279, Morristown, NJ, USA, 1989. Association for Computational Linguistics.
- [9] J. Joseph J. LaViola. A survey of hand posture and gesture recognition techniques and technology. Technical report, Brown University, Providence, RI, USA, 1999.
- [10] J. José Josemar de Oliveira, J. M. de Carvalho, C. O. de A. Freitas, and R. Sabourin. Evaluating nn and hmm classifiers for handwritten word recognition. In *SIBGRAPI '02: Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing*, pages 210–217, Washington, DC, USA, 2002. IEEE Computer Society.
- [11] N. Liu and B. C. Lovell. Gesture classification using hidden markov models and viterbi path counting. In *DICTA*, pages 273–282, 2003.
- [12] J. A. Montero and L. E. Sucar. Feature selection for visual gesture recognition using hidden markov models. *enc*, 0:196–203, 2004.
- [13] M. Morita and L. S. Oliveira. Introdução aos modelos escondidos de markov. Technical report, PPGIA-PUCPR, Curitiba-Brazil, November 1998.
- [14] H. Pistori, P. S. Martins, M. C. Pereira, and J. J. Neto. Sigus - plataforma de apoio ao desenvolvimento de sistemas para inclusão digital de pessoas com necessidades especiais. *IV Congresso Iberdiscap: Tecnologias de Apoio a Portadores de Deficiência, Vitória*, February 2006.
- [15] D. Prescher. A tutorial on the expectation-maximization algorithm including maximum-likelihood estimation and em training of probabilistic context-free grammars. *ArXiv Computer Science e-prints*, December 2004.
- [16] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *IEEE Computer Graphics and Applications*, 77(2):42–53, 1989.
- [17] F. C. C. W. D. Raphael. *Dicionário Enciclopédico Ilustrado Trilíngüe da Língua de Sinais Brasileira*, volume Volume I: Sinais de A a L. São Paulo, SP: Edusp, Imprensa Oficial, 2002.
- [18] F. C. C. W. D. Raphael. *Dicionário Enciclopédico Ilustrado Trilíngüe da Língua de Sinais Brasileira*, volume Volume I: Sinais de M a Z. São Paulo, SP: Edusp, Imprensa Oficial, 2002.
- [19] K. Siriboon, A. Jirayusakul, and B. Kruatrachue. Hmm topology selection for on-line thai handwritten recognition. In *CW '02: Proceedings of the First International Symposium on Cyber Worlds (CW'02)*, page 0142, Washington, DC, USA, 2002. IEEE Computer Society.
- [20] C. A. Ynoguti. *Reconhecimento de Fala Contínua Usando Modelos Ocultos de Markov*. PhD thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, São Paulo, Brasil, 1999.
- [21] L.-G. Zhang, X. Chen, C. Wang, Y. Chen, and W. Gao. Recognition of sign language subwords based on boosted hidden markov models. In *ICMI '05: Proceedings of the 7th international conference on Multimodal interfaces*, pages 282–287, New York, NY, USA, 2005. ACM Press.
- [22] M. Zimmermann and H. Bunke. Hidden markov model length optimization for handwriting recognition systems. In *IWFHR '02: Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02)*, page 369, Washington, DC, USA, 2002. IEEE Computer Society.