

Um Modelo de Máquinas de Vetores de Suporte Estruturadas em Árvore Binária Probabilística Aplicado ao Reconhecimento de Posturas Manuais em Tempo Real

Ticiano A. C. Bragatto, Gabriel I. S. Ruas, Augusto P. Martins e Marcus V. Lamar

Resumo — Um novo modelo de classificador baseado em Máquinas de Vetores de Suporte (SVM) estruturadas em árvore binária é proposto e aplicado ao reconhecimento visual de posturas manuais da Linguagem Americana de Sinais. A nova estrutura de árvore é baseada no processo de geração do código de Huffman, visando diminuir a complexidade computacional sem o comprometimento da taxa de acerto. O novo método é comparado com técnicas clássicas de reconhecimento de padrões, tendo a distribuição de probabilidades da taxa de reconhecimento e a complexidade computacional como critérios de avaliação. Os resultados mostram que o modelo probabilístico de SVM proposto é mais rápido e atinge maiores taxas médias de acerto entre os métodos testados, em uma aplicação de reconhecimento de posturas manuais em tempo real.

Palavras-Chave — Reconhecimento de Gestos, Máquinas de Vetores de Suporte

Abstract — This paper proposes the use of a probability optimized binary tree structured multi-class SVM for real-time visual based American Sign Language finger spelling recognition. The structure of the binary tree is defined as a Huffman code tree, based on the probability of occurrence of each letter in American English. This novel method is then compared with well known classical pattern recognition methods, with computational effort and recognition rates as criteria. The results show that probability optimized binary tree SVM is more accurate and faster than other classical methods tested.

Key Words — Gesture recognition, Support Vector Machines

I. INTRODUÇÃO

Atualmente, a comunicação entre homens e máquinas é feita principalmente por meio do uso de dispositivos como teclado e mouse. Pesquisas visando ao desenvolvimento de técnicas que permitam aos computadores ver e ouvir são muito importantes para a criação de um sistema capaz de interagir de forma natural com os seres humanos. O reconhecimento de voz é um campo de estudos bastante popular e muitos trabalhos têm sido publicados ao longo das últimas décadas.

Com o incremento do poder computacional dos computadores pessoais, o reconhecimento de gestos tem atraído a atenção de vários pesquisadores. O desenvolvimento de sistemas de processamento de vídeo de alta complexidade em tempo real ainda é um problema bastante desafiador, atingindo facilmente os limites de poder de processamento dos modernos computadores. Um dos mais complexos e gramaticalmente estruturados conjuntos de gestos humanos é o associado à linguagem dos sinais utilizada pelas pessoas com deficiência auditiva, que usam um vocabulário de mais de 8000 diferentes gestos. No contexto da linguagem dos sinais, a soletração de palavras é uma importante ferramenta para descrever nomes de pessoas, lugares e designar novos objetos.

Na área de reconhecimento de padrões, Redes Neurais Artificiais e Máquinas de Vetores de Suporte (SVM – *Support Vector Machines*) têm se mostrado como técnicas concorrentes em diversas aplicações. Para a implementação de um sistema de reconhecimento em tempo real, além da qualidade do classificador, sua complexidade computacional é um fator muito importante a ser considerado. Neste trabalho propõe-se um novo classificador SVM multiclasse estruturado em árvore, o qual é projetado para minimizar sua complexidade computacional de acordo com a probabilidade de ocorrência de cada classe, chamado aqui SVM estruturada em Árvore Binária Otimizada pela Probabilidade (*Probability Optimized Binary Tree structured Support Vector Machine - POBT-SVM*). O desempenho do novo classificador e sua complexidade computacional são comparados experimentalmente com Redes Neurais Perceptron Multicamadas (MLP – *Multi Layer Perceptron*), com outras estruturas propostas na literatura de classificadores multiclasse baseadas em SVM, e com os classificadores clássicos de *Template Matching* (TM) e *k-Nearest Neighbors* (k-NN).

II. SISTEMA DE RECONHECIMENTO DE POSTURAS MANUAIS

O diagrama geral do sistema de reconhecimento de posturas manuais usado neste trabalho é apresentado na Fig. 1. Este sistema completo visa detectar, analisar e classificar posturas manuais de forma a reconhecer a palavra ou frase dita para posterior vocalização.

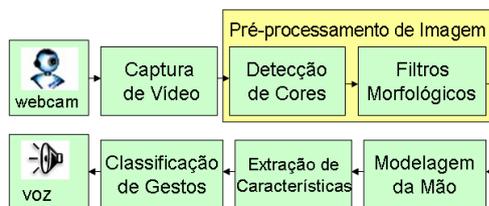


Figura 1. Diagrama em blocos do sistema de visão computacional

O sistema é composto por um estágio de captura de vídeo, que trabalha a uma taxa máxima de 29.97 quadros por segundo (qps), correspondente às atuais webcams. Cada quadro é analisado pelo estágio de pré-processamento, que executa a detecção da mão na imagem. Este trabalho está focado na classificação das características extraídas da imagem da mão já segmentada e modelada. Esta etapa de pré-processamento utiliza-se da energia da imagem obtida da diferença entre quadros consecutivos para avaliar se o gesto ainda está sendo executado. O segmentador irá atuar apenas quando este estágio detectar uma imagem quase estática, uma vez que neste trabalho estamos interessados apenas na análise de posturas da mão correspondente à soletração usando o alfabeto manual da Linguagem Americana de Sinais (ASL).

A fim de reduzir o esforço computacional necessário à etapa de localização da mão e seu rastreamento, bem como melhorar a precisão do vetor de características extraído do sinal de vídeo, uma luva colorida é usada, em que cada dedo possui uma cor diferente, conforme a Fig. 2. Esta metodologia visa a reduzir a complexidade computacional do sistema. Técnicas de detecção e extração de características de posturas manuais usando imagem da mão livre, ou seja, sem o uso de luvas, estão sendo propostas. Porém estas técnicas apresentam alta complexidade computacional, não sendo adequadas para aplicações em tempo real com os recursos computacionais disponíveis hoje em dia.

O uso de uma luva colorida tem sido adotado em diversas pesquisas [2][3][4][5] devido a sua eficiência para a redução da complexidade computacional da etapa de pré-processamento. O uso de luvas instrumentalizadas de alto custo [6], ou mesmo as recentes e mais baratas luvas baseadas em acelerômetros [7], torna necessário o uso de equipamentos e circuitos específicos para a aquisição do sinal. Uma simples luva colorida permite rápida detecção da mão, rastreamento e modelagem precisos da posição e posturas dos dedos através de uma *webcam*. Sistemas que operam a mão livre necessitam de uma complexa etapa de detecção de cor de pele, além de estágios de modelagem bastante elaborados [4][8][9]. Isto que pode impedir o desenvolvimento de sistemas que operem em tempo real com os recursos computacionais ainda limitados oferecidos pelos modernos computadores portáteis.

O uso de luvas coloridas permite que o sistema seja simplificado, sendo composto basicamente de um reconhecedor de cores – para as seis cores da luva – que resulta em uma série de imagens binárias, uma para cada cor. Com a utilização de filtros morfológicos de erosão e dilatação as regiões de interesse são detectadas.

Após o estágio de detecção e pré-processamento, um

vetor de características de dimensão 20 é extraído da imagem de acordo com o modelo proposto por [10]. O presente trabalho está focado nas técnicas de classificação deste vetor de características. O desempenho será avaliado baseado apenas nas medidas da taxa de acerto e complexidade computacional exigida nesta etapa.



Figura 2. Luva colorida e ambiente de trabalho

III. MÁQUINAS DE VETORES DE SUPORTE ESTRUTURADAS EM ÁRVORE BINÁRIA OTIMIZADA PELA PROBABILIDADE

As máquinas de vetores de suporte (SVM – *Support Vector Machines*) são um conjunto de métodos de aprendizagem supervisionada usados principalmente em problemas de classificação e regressão [11]. A idéia principal das SVM é a maximização da margem de separação entre duas classes, buscando a minimização da taxa de erro de classificação. Técnicas como a *soft margin* e *kernel trick* podem ser utilizadas para solucionar problemas mais complexos [12].

O uso de SVM para classificação resulta em uma saída binária de valor 1 ou -1, de acordo com a posição do vetor de entrada em relação ao hiperplano separador. Para resolver problemas multiclases com SVMs [13], têm sido propostas várias técnicas, sendo as mais populares o um-contra-um (OvO - *one-versus-one*) e um-contra-resto (OvR - *one-versus-rest*). A primeira (OvO SVM) possui alta complexidade computacional devido à necessidade de avaliação de $k(k-1)/2$ SVMs para um problema com k classes. A segunda (OvR SVM) é composta de apenas $(k-1)$ SVMs, sendo, portanto, de mais rápida avaliação. No entanto, o problema torna-se de mais difícil solução, implicando em uma etapa de treinamento mais onerosa, e gerando, na maioria das vezes, aumento da taxa de erros do sistema.

O esforço computacional exigido pela técnica OvR SVM na classificação cresce linearmente com o número de classes k . Uma possível forma de reduzir ainda mais esta complexidade é o uso de classificadores estruturados hierarquicamente em árvore binária. Diversas propostas de estruturas em árvores de classificadores baseados em SVM têm sido apresentadas [14][17]. O principal interesse nestes estudos é a definição da estrutura da árvore binária que maximiza a taxa de acerto do classificador. No presente trabalho, visamos desenvolver técnicas de redução da complexidade computacional do classificador. A estrutura da árvore é organizada de forma a posicionar as classes mais prováveis nos nós próximos à raiz. Isto tende a reduzir o esforço computacional médio exigido

pele classificador. Este trabalho tem por objetivo o reconhecimento da soletração manual usada na Linguagem Americana de Sinais (ASL). A Tabela I apresenta a probabilidade de ocorrências das letras no idioma inglês, de forma que podemos propor uma estruturação simples em árvore binária progressiva (PT-SVM – *Progressive Tree SVM*), apresentada na Fig. 3. Deve-se notar que à medida que avançamos na estrutura da árvore, os classificadores SVM dos nós filhos vão tornando-se menos complexos, uma vez que esses classificadores devem confrontar uma letra com o conjunto do restante, reduzindo ainda mais a complexidade computacional do método.

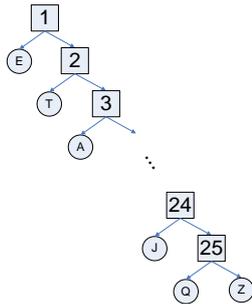


Figura 3. SVM estruturada em árvore binária progressiva para reconhecimento de letras na ASL.

Este trabalho propõe e testa uma estrutura chamada SVM estruturada em Árvore Binária Otimizada pela Probabilidade (POBT-SVM – *Probability Optimized Binary Tree Structured Support Vector Machines*). Esta nova estrutura visa a reduzir a complexidade computacional nos estágios de treino e uso, sem ocasionar perdas na qualidade do classificador.

A POBT-SVM é projetada similarmente às SVM em árvore binária [17]. Porém, com o objetivo de minimizar o esforço computacional, a árvore binária é construída de acordo com o *Entropy-Constrained Huffman Coding* [15], baseando-se na probabilidade de ocorrência de cada classe encontrada nos problemas reais.

A Fig.4 apresenta a estrutura da árvore final projetada de acordo com a codificação de Huffman obtida para a frequência de ocorrência das letras no idioma inglês [16]. Os nós (quadrados) são classificadores SVM, e as folhas (círculos) as classes correspondentes às letras. O vetor de entrada a ser classificado é apresentado inicialmente ao classificador correspondente à raiz da árvore, SVM do nó número 25 na Fig. 4. Para cada resultado de valor 1 de saída SVM é seguido o caminho da direita na árvore, e para cada resultado de valor -1 é seguida a direção esquerda. O processo continua até chegar a um nó folha, que define então a classe do vetor de entrada.

Esta nova estrutura minimiza o esforço computacional global do problema de classificação de acordo com a probabilidade de ocorrência de cada classe em relação ao número de classificadores utilizados. No entanto, a complexidade de cada classificador depende do mapeamento não linear utilizado pelo SVM e da quantidade de vetores de suporte necessários. Nos testes a seguir é apresentada esta análise.

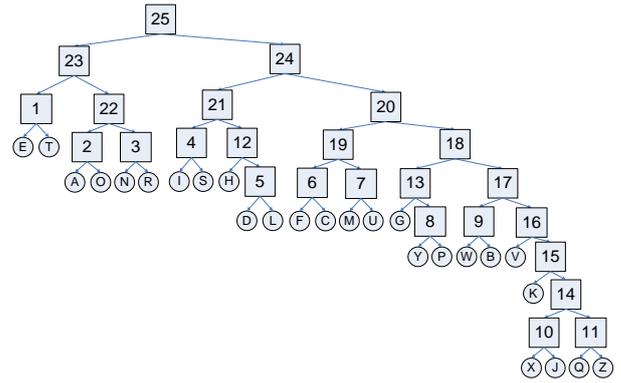


Figura 4. POBT-SVM para a probabilidade de ocorrência das letras no idioma inglês.

IV. RESULTADOS EXPERIMENTAIS

Os resultados apresentados nesta seção foram obtidos usando um computador Pentium IV 2.4GHz, 512 Mbytes RAM e uma webcam com captura de quadros a 24 bits/pixel, 352x288 pixels a 29,97 quadros por segundo. O ambiente de teste é de um típico laboratório de pesquisa, com fundo complexo e sem controle especial de iluminação, conforme apresentado nas Fig. 2 e 5.

O conjunto de 26 posturas manuais usado nos experimentos é mostrado na Fig. 5. Para o projeto dos classificadores, os 1560 vetores amostrados (60 amostras de cada postura manual) foram divididos em 50% para o conjunto de treino e 50% para o conjunto de teste.

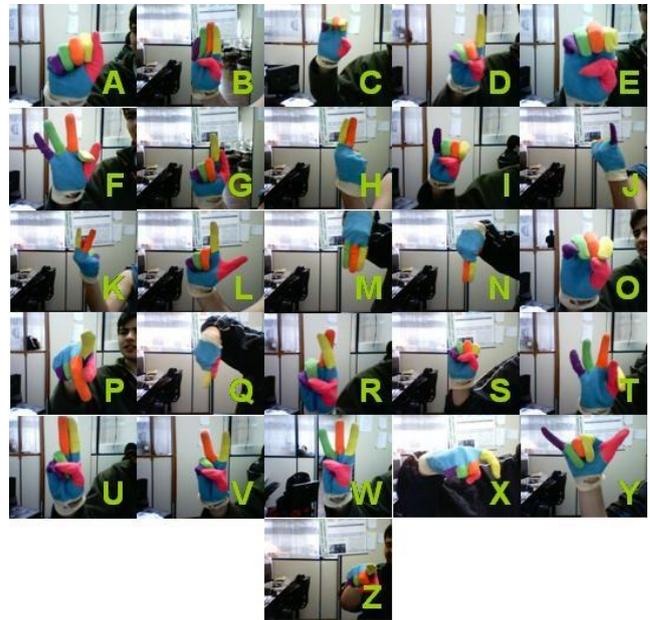


Figura 5. Posturas manuais da ASL para soletração

Para classificar as 26 posturas, 25 SVMs foram treinadas usando kernels polinomiais, devido ao seu reduzido esforço computacional quando comparados à utilização de kernel gaussiano ou outros mais complexos.

A Tabela I apresenta os resultados para a árvore binária progressiva (PT-SVM), ordenados de acordo com a probabilidade de ocorrência da classe (letra) no idioma inglês americano.

TABELA I: ICC PARA A PT-SVM

Letra	Probabilidade	ICC do SVM	ICC da Letra	Contribuição
E	12.70%	1724	1724	218.98
T	9.06%	1636	3361	304.37
A	8.17%	2152	5514	450.33
O	7.51%	3098	8613	646.58
N	6.97%	1514	10128	705.52
R	6.75%	1724	11853	799.96
I	6.33%	3141	14995	948.73
S	6.09%	2453	17449	1063.34
H	5.99%	617	18067	1081.67
D	4.25%	1560	19628	834.78
L	4.03%	1507	21136	850.72
F	2.78%	1109	22246	618.88
C	2.76%	1894	24141	665.81
M	2.41%	1052	25194	606.17
U	2.36%	1304	26499	625.38
G	2.23%	1109	27609	615.13
Y	2.02%	740	28350	571.25
P	1.97%	590	28941	571.30
W	1.93%	632	29574	570.48
B	1.49%	412	29987	447.41
V	0.98%	658	30646	299.72
K	0.77%	590	31237	241.15
X	0.15%	658	31896	48.80
J	0.15%	968	32865	49.30
Q	0.10%	289	33155	31.50
Z	0.07%		33155	24.54

Na Tabela I, um Índice de Complexidade Computacional (ICC) é associado a cada SVM e calculado com base no número e no tipo de operações matemáticas necessárias à sua avaliação seguindo-se a metodologia proposta por [18]. Uma vez que o classificador é organizado hierarquicamente, o ICC de uma letra pode ser calculado como a soma dos ICCs das SVMs das letras anteriores na árvore binária, de acordo com

$$ICC_n = 1 + \sum_{k=1}^n ICC_{SVM_k} \quad (1)$$

onde ICC_{SVM_k} é o índice de complexidade da avaliação da k -ésima SVM, e ICC_n corresponde ao esforço computacional para classificar um vetor de entrada na classe n , chamado aqui ICC da letra. As diferenças entre os ICC_{SVM} se devem ao número de vetores de suportes do classificador.

O SVM mais complexo é o correspondente à letra 'I', que necessita 73 vetores de suporte e um kernel polinomial de 3º grau. Entretanto, para as letras menos prováveis ('Q' e 'Z'), o sistema deve executar toda a árvore, resultando em ICC_{26} de 33155.

O ICC global do sistema pode ser estimado pela soma ponderada de cada ICC da letra e sua probabilidade de ocorrência, como

$$ICC_{PT-SVM} = \sum_{n=1}^{26} ICC_n \times P(n) \quad (2)$$

onde ICC_n é o esforço computacional para a classificação da n -ésima letra e $P(n)$ sua probabilidade de ocorrência. O ICC calculado pela Eq. (2) resulta em um valor médio de 13892 para o reconhecimento de letras soletradas em ASL.

A Tabela II apresenta o SVM correspondente a cada nó da árvore na topologia POBT-SVM, sua probabilidade de

avaliação, o ICC necessário para sua computação e o quanto corresponde sua contribuição para o ICC total do sistema, considerando as probabilidades de ocorrência já apresentadas na Tabela I.

TABELA II: ICC PARA A POBT-SVM

Número do SVM	Probabilidade	ICC do SVM	Contribuição
1	21.76%	494	107.48
2	15.67%	716	112.23
3	13.72%	412	56.51
4	12.42%	658	81.73
5	8.28%	576	47.68
6	5.54%	535	29.64
7	4.77%	453	21.59
8	3.99%	535	21.34
9	3.42%	412	14.09
10	0.30%	535	1.62
11	0.17%	674	1.14
12	14.27%	453	64.62
13	6.22%	740	46.01
14	0.47%	464	2.19
15	1.24%	371	4.62
16	2.22%	453	10.07
17	5.64%	371	20.94
18	11.86%	576	68.31
19	10.31%	535	55.14
20	22.17%	576	127.68
21	26.69%	412	109.95
22	29.39%	576	169.28
23	51.15%	494	252.67
24	48.85%	494	241.33
25	100.00%	632	632.00

O SVM mais complexo é o de número 13, que na Fig. 4 distingue a letra 'G' das letras 'Y' e 'P'. Este SVM é composto de 18 vetores de suporte. O pior caso na POBT-SVM para a soletração em ASL também ocorre com as letras menos prováveis ('Q' e 'Z'). No entanto, para esta avaliação, é necessário a classificação por apenas 9 SVMs (nós 25, 24, 20, 18, 17, 16, 15, 14 e 11), resultando em um ICC de 4611, muito menor que o valor 33155 obtido na estrutura anterior.

O ICC global do sistema pode ser calculado por

$$ICC_{POBT-SVM} = \sum_{n=1}^{25} ICC_n \times P_n \quad (3)$$

onde ICC_n é o esforço computacional da n -ésima SVM e P_n sua probabilidade de avaliação. Esta probabilidade é calculada como a soma das probabilidades de ocorrência de cada letra que se encontra em nós inferiores ao nó do SVM avaliado. O ICC resultante do sistema é estimado em 2300.

Apresentamos a seguir uma comparação de uso da técnica proposta (POBT-SVM), o SVM multiclasse no formato um-contra-um (OvO SVM), um-contra-resto estruturado em árvore (PT-SVM), Rede Neural MultiLayer Perceptron (MLP), k vizinhos mais próximos (k -NN, k -Nearest Neighbors) e Casamento de Padrão (TM, *Template Matching*).

Para o k -NN e TM, foram testadas as distâncias Euclidiana e *City Block*, devido aos seus diferentes requisitos computacionais e qualidade de reconhecimento. Para o k -NN, o valor $k=1$ (Vizinho Mais Próximo) apresentou o melhor resultado, com a variação de k no

conjunto $\{1, 3, 5, \dots, 25\}$. O vetor padrão na técnica TM foi calculado como o vetor centróide (média aritmética) de cada classe do conjunto de treino.

A rede neural treinada é uma *Multi-Layer Perceptron* (MLP) de 3 camadas, função de ativação sigmoide, contendo 20 entradas, 26 unidades escondidas e 26 unidades de saída, treinada com o algoritmo de *Backpropagation* com 3.000 épocas e taxa de aprendizagem de 0,1.

A SVM um-contra-um (OvO SVM) trabalha como um sistema de votação. Para cada postura manual, correspondente a uma letra, 25 SVMs são treinados. No estágio de classificação, o vetor de teste é apresentado a todos os $26 \times 25 / 2 = 325$ SVMs treinados. Cada classe possui, então, um determinado número de votos dados pelas SVMs. A classe mais votada é considerada a saída do classificador. Há dois problemas principais no

classificador OvO SVM: a elevada complexidade computacional necessária à avaliação de todas as SVMs e a possibilidade de um empate no sistema de votação.

Dada a característica estatística dos treinamentos de classificadores baseados em amostras, a fim de avaliar e comparar o desempenho dos classificadores de maneira justa foi realizado uma série de treinos e testes. Conjuntos disjuntos de treino e de teste são aleatoriamente sorteados sem repetição de vetores de amostras. A cada conjunto um classificador é treinado, testado e seu desempenho é computado em termos de taxa de acerto. A Fig. 6 apresenta os resultados para lotes de 10.000 treinamentos de cada classificador, variando os conjuntos de treino e teste, mantendo a proporção de 50% dos vetores para treino e 50% para testes de cada classe.

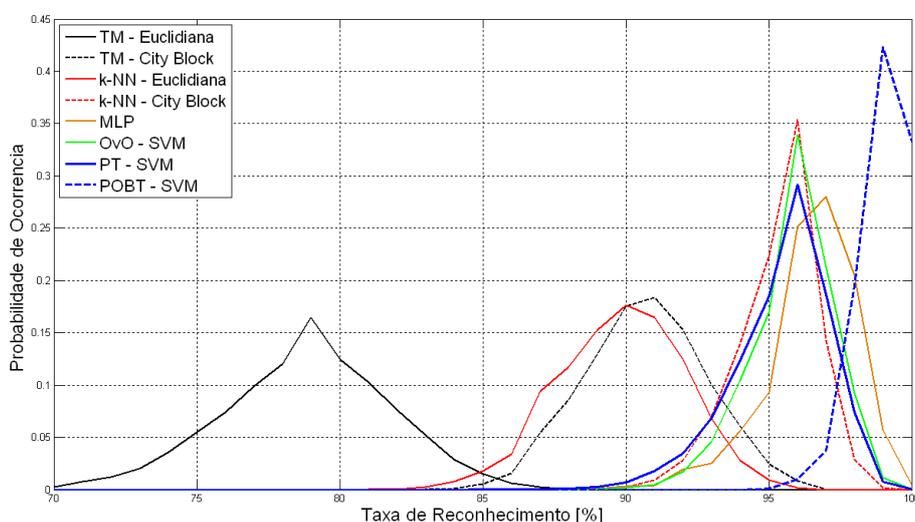


Figura 6. Taxa de reconhecimento versus Probabilidade de Ocorrência

A Fig. 6 apresenta a função densidade de probabilidade (*pdf*) da taxa de correto reconhecimento para os classificadores testados. Pode-se notar, por esta figura, que o classificador TM com distância Euclidiana apresenta baixa exatidão, por apresentar baixa taxa média de reconhecimento, em torno de 78,8%. Sua baixa precisão indica uma maior sensibilidade às amostras de treino, que pode ser verificada pelo maior espalhamento (desvio padrão) da sua *pdf*. Por outro lado, os classificadores k-NN com distância *City Block*, OvO-SVM e PT-SVM apresentam desempenho similar em termos de taxa de acerto esperada 95,5% e menor sensibilidade aos diferentes conjuntos de vetores de treino. A rede neural MLP apresenta-se como um classificador mais eficiente em termos de exatidão, porém com uma precisão um pouco menor que pode ser notado pelo seu maior espalhamento. O classificador proposto POBT-SVM apresenta a maior taxa esperada, de 99,0%, e o menor desvio padrão, 0,78%, indicando sua melhor precisão e robustez quanto à divisão das amostras nos conjuntos de treino e teste.

TABELA III: RESULTADOS DA CLASSIFICAÇÃO DE POSTRAS MANUAIS

Classificador	ICC Esperado	Taxa Esperada	Desvio Padrão
TM – Euclidean	1,560	78,8%	2,98%
TM – City Block	1,050	90,6%	2,08%
K-NN – Euclidean	31,210	89,7%	2,17%
K-NN – City Block	21,018	95,3%	1,46%
MLP	4,182	96,5%	1,66%
OvO SVM	155,449	95,8%	1,50
PT-SVM	13,892	95,4%	1,77%
POBT-SVM	2,300	99,0%	0,78%

A Tabela III apresenta resultados numéricos obtidos em termos de complexidade computacional média, taxa de acerto esperada e desvio padrão da taxa de reconhecimento dos classificadores testados, cujas *pdf* são mostradas na Fig. 6.

Neste experimento, a POBT-SVM superou em qualidade os outros métodos testados, atingindo 99,0% de taxa média de reconhecimento, com ICC de 2.300, sendo mais complexo apenas que as técnicas de TM. O TM com distância *City Block* é o classificador menos complexo computacionalmente, com ICC de 1.050. No entanto, apresenta uma taxa de reconhecimento esperada de 90,6%.

A elevada complexidade computacional do OvO SVM se deve à necessidade de avaliação de todos os 325 SVMs. Por outro lado, PT-SVM necessita calcular no pior caso 25 SVMs, enquanto POBT-SVM necessita no máximo 9 avaliações, sendo, portanto, mais adequado em aplicações de tempo real.

É de senso comum que o tempo necessário de processamento é uma medida consistente da complexidade computacional de um algoritmo. Para testar o impacto dos classificadores em um experimento em tempo real, foi pedido que um usuário soletrasse vários nomes próprios, de pessoas, lugares e objetos, escolhidos de modo a obedecer à probabilidade de ocorrência das letras no idioma inglês.

Os classificadores k-NN, OvO SVM e PT-SVM não foram capazes de operar em tempo real, devido a suas elevadas complexidades computacionais. O classificador OvO SVM atingiu a taxa de 1 qps e PT-SVM 8,5 qps. Usando MLP, o sistema atingiu 28,1. Entretanto, usando POBT-SVM, foi atingido o limite imposto pelo sistema de aquisição de vídeo de 29,97 qps.

V. CONCLUSÕES

Este trabalho propôs uma nova estrutura para classificadores baseados em máquinas de vetores de suporte, denominada POBT-SVM. A estrutura de árvore binária otimizada pela probabilidade de ocorrência de cada classe projetada pelo processo de codificação de Huffman mostrou-se adequada à estruturação hierárquica de SVMs em aplicações em tempo real, em que um baixo custo computacional do classificador é necessário, sem aumentar a taxa de erros de classificação. Comparações com técnicas de classificação clássicas, como redes neurais e k-NN, foram feitas em um experimento de reconhecimento visual da soletração de palavras usando as posturas manuais da Linguagem Americana dos Sinais.

O uso de redes neurais para a classificação das posturas manuais permite uma redução do tempo de processamento do sistema, quando comparado com o classificador OvO SVM, atingindo taxas de 28,1 qps para sistemas de reconhecimento contínuo baseado em sinais de vídeo da soletração manual, sem perdas perceptíveis de qualidade do reconhecimento; 96,5% realizado pela MLP, enquanto OvO SVM atinge 95,8% de taxa média de reconhecimento a uma taxa de quadros de 1 qps.

A técnica proposta POBT-SVM mostrou ser mais rápida computacionalmente que MLP e com superior qualidade de reconhecimento. No experimento, reduziu a complexidade computacional aumentando a taxa de quadros de 28,1 qps para 29,97qps, que é o limite imposto pelo sistema de aquisição de vídeo utilizado, atingindo 99,0% de taxa esperada de reconhecimento.

Trabalhos futuros incluem estudos para integrar o treinamento das SVM ao projeto da estrutura da árvore binária, de forma a desenvolver uma metodologia de treinamento de classificadores SVM estruturados em árvore binária que sejam de complexidade computacional mínima, dada uma distribuição de probabilidade de classes esperada e SVM com kernels polinomiais. Estudos utilizando as posturas do alfabeto manual da Língua

Brasileira dos Sinais (LIBRAS) e a probabilidade do Português brasileiro também serão realizados.

REFERÊNCIAS

- [1] S. Tamura and S. Kawasaki, "Recognition of sign language Motion images", in *Pattern Recognition*, Vol.21, No 4, 1988, pp. 343-353.
- [2] T. Stamer, and A. Pentland, "Visual recognition of american sign language using hidden markov models". *Int. Workshop on Aut. Face and Gesture Recognition*, 1995.
- [3] Y. Iwai, K. Watanabe, Y. Yagi, and M. Yachida, "Gesture recognition using colored gloves". *IEEE Int. Conf. on Pattern Recognition*, 1996, pp.662-666.
- [4] K. Abe and S. Ozawa, "3-D Drawing system via hand motion recognition from two cameras", *IEEE Int Conf. on Systems, Man, and Cybernetics*, v. 2, 2000, pp. 840-845
- [5] M. Bray, H. Sidenbladh and J.-O. Eklundh, "Recognition of gestures in the context of speech", *Proc. of the 16th Int. Conf. on Pattern Recognition*, v. 1, 2002, pp. 356-359
- [6] R.-H. Liang, M. Ouhyoung, "A real-time continuous gesture recognition System for sign language", *Proc. of IEEE 3rd. Int. Conf. on Face and Gesture Recognition*, 1998, pp. 558
- [7] J. L. Hernandez-Rebollar, N. Kyriakopoulos, and R. W. Lindeman, "A whole hand interface for virtual reality", *ACM SIGGRAPH Conf. Abstracts and App.*, 2002, pp. 259.
- [8] S. G. Wysocki, M. V. Lamar, S. Kuroyanagi, and A. Iwata, "A rotation invariant approach on static-gesture recognition using boundary histograms and neural networks.", *9th Int. Conf. on Neural Information Processing*, 2002, pp.726- 729.
- [9] Vassalis, and S. Stan, "Estimating 3D hand pose from a cluttered image", *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2003, pp. 432-439.
- [10] M. V. Lamar, M. S. Bhuiyan, A. Iwata, "Hand gesture recognition using T-CombNET: A new neural network model", *IEICE Trans. on Information and Systems*, Vol. E83-D, N.11, 2000, pp. 1986-1995.
- [11] V. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, 1998.
- [12] B. E. Boser, I. M. Guyon, and V. N. Vapnik. "A training algorithm for optimal margin classifiers." Haussler, editor, *5th Annual ACM Workshop on COLT*, 144-152, Pittsburgh, PA, 1992.
- [13] J. Weston, and C. Watkins, "Multi-class support vector machines". *Technical Report CSD-TR-98-04*, 1998.
- [14] B. Fei and J. Liu, "Binary tree of SVM: a new fast multiclass training and classification algorithm", *IEEE Trans on Neural Networks*, Vol.17, Iss.3, pp. 696- 704,2006.
- [15] K. Sayood, *Introduction to Data Compression*, Morgan Kaufmann Publishers, 2000.
- [16] R. Lewand, *Cryptological Mathematics*, The Mathematical Association of America, 2007.
- [17] S. Cheong, S. H. Oh, and S.-Y. Lee, "Support vector machines with binary tree architecture for multi-class classification", *Neural Info. Process.—Lett. Rev.*, vol. 2, no. 3, Mar. 2004.
- [18] T.A.C. Bragatto, V.A.P. Benso, J. Sugawara, J. ; M.V. Lamar, "Reconhecimento de gestos em Tempo-Real Utilizando uma Rede Neural Artificial de Baixa Complexidade Computacional Para Detecção de Cores". *Proc.of XVIII Brazilian Symp. on Computer Graphics and Image Processing*, 2005.