Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, Kobe International Conference Center, Kobe, Japan, December 15-17,

Acoustic Feature Selection Utilizing Multiple Kernel Learning for Classification of Children with Autism Spectrum and Typically Developing Children

Yasuhiro Kakihara, Tetsuya Takiguchi, Yasuo Ariki, Yasushi Nakai and Satoshi Takada

Abstract— **This paper reports the result of a classification experiment carried out using acoustic features for children with autism spectrum, where a new feature-weighting method using a multiple kernel learning (MKL) algorithm is proposed for classification between children with autism spectrum and typically developing children. Our MKL-SVM simultaneously estimates both the classification boundary and weight of each acoustic feature, where 484 acoustic features are used in our experiments. The estimated weight indicates how acoustic features are useful for classification. Our results show the large weight acoustic features mainly for line spectral frequencies in the classification experiment using acoustic features for children with autism spectrum.**

I. INTRODUCTION

Many speech recognition technologies have been studied for adults to date. Recently, research has been carried out for children, elderly people or people with disabilities [1], [2], and research related to autistic spectrum disorders (ASD) has also been focused on. An autistic spectrum obstacle is a congenital cerebral dysfunction, and it is a type of the developmental disease causing difficulty in communication, perceptual, cognitive, and linguistic functions [3], [4], [5]. Children with autistic spectrum are diagnosed as having less than normal social interaction and linguistic skills, and have restricted interests and a stereotypic pattern of behavior [6], [7], [8]. It is estimated that the rate of autistic disorders, such as Asperger disorder [9] and nonspecific pervasive developmental disorders, is from 1% to 2% of all children [10]. Since the symptoms of an autistic spectrum obstacle are the result of a variety of causes, a fundamental, all-encompassing medical treatment is difficult. However, for an autistic spectrum obstacle, early detection and suitable education can have a significant impact on future social prognosis. Recent investigations have demonstrated that the early support which specialized in autistic spectrum obstacle (such as Picture Exchange Communication System (PECS) [11], Applied Behavier Analysis (ABA) [12], Social Skills Training (SST) [13]) is effective [7]. In the field of acoustic technology, however, there has been little research focused on discriminating between children with autism spectrum and typically developing children.

This paper reports the result of a classification experiment carried out using acoustic features for children with autism spectrum, where a new weighting method using a multiple kernel learning (MKL) algorithm is proposed. Our MKL-SVM simultaneously estimates both the classification boundary and weight of each acoustic feature (484 acoustic features were used in our experiments). The estimated weight indicates how acoustic features are useful for the intended classification.

A multiple kernel learning (MKL) algorithm is a machine learning-based technique for learning the proper weights of the corresponding kernels, which uses multiple classifiers with a kernel. MKL has been used as an integration method by calculating the appropriate weights corresponding to each kernel, while classical kernel-based methods (such as Support Vector Machines) are based on a single kernel only. In the field of image processing research, object recognition methods based on MKL have been proposed for integrating image features [14].

In our experiments, we recorded the speech data of children. The children ranged in age from kindergarteners to the fourth graders. This database consists of 20 ASD children and 21 typically developing children. From this database, we extracted 484 acoustic features,

Y. Kakihara, T. Takiguchi and Y. Ariki are with Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada, Kobe, Hyogo 657-8501, Japan kakyhara@me.cs.scitec.kobe-u.ac.jp, [takigu,ariki]@kobe-u.ac.jp

Y. Nakai is with Kawasaki College of Allied Health Professions, 316 Matsushima, Kurashiki, Okayama 701-0194, Japan nakai@jc.kawasaki-m.ac.jp

S. Takada is with Graduate School of Health Science, Kobe University, 7-10-2 Tomogaoka, Suma, Kobe, Hyogo 654-0142, Japan satoshi@kobe-u.ac.jp

978-1-4799-2625-1/13/$31.00 ©2013 IEEE

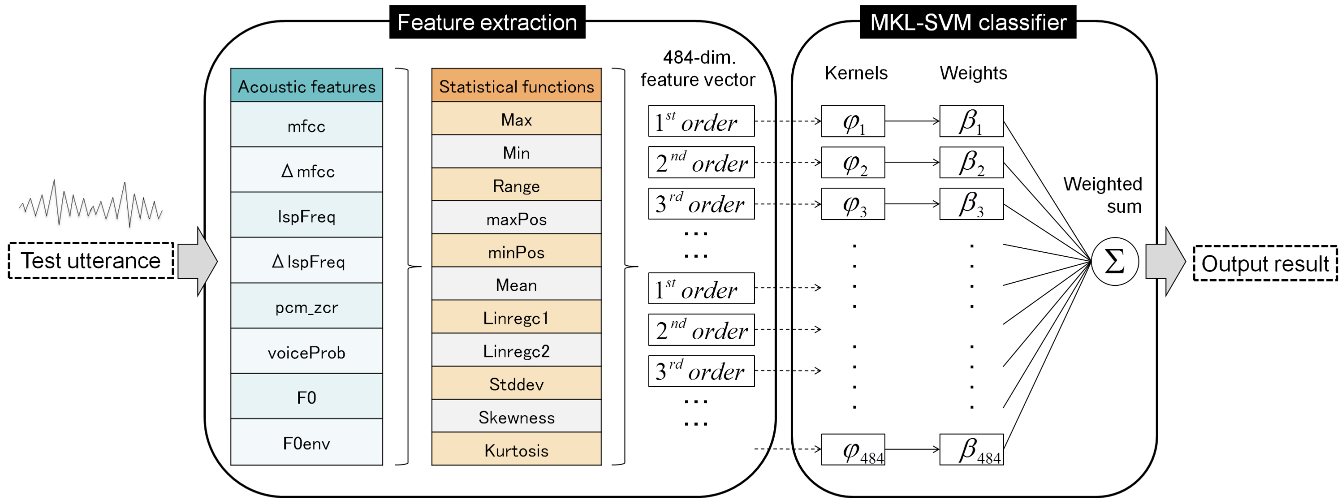

Fig. 1. System outline for MKL-SVM

such as mel-frequency cepstral coefficients, linear predictive coding (LPC) coefficients, and so on. Fig. 1 shows an outline of our MKL-SVM. MKL-SVM was performed for selection and weighting for acoustic features for discriminating between children with autism spectrum and children who are developing typically. In our MKL-based approach, 484 acoustic features are separately weighted. Each weight value calculated by MKL indicates how useful each acoustic feature is for classifying the speech data of children with autism spectrum and typically developing children.

II. Feature Selection and Classification Using MKL-SVM

In this paper, we used an open-source openEAR [15], [16] to extract features, such signal energy, FFT-spectrum, mel-spectrum, MFCC, line spectral frequencies (line spectral pairs), pitch, voice quality (Harmonics-To-Noise Ratio), LPC coefficients, Perpetual Linear Predictive coefficients, formants, and so on. In our experiments, 484 acoustic features are used for classification of children with autism spectrum and typically developing children. We extracted 484 dimensional features per word by using openEAR. The 484 dimensional features consist of combination 44 acoustic features (such as formant feature, stress accent, pitch accent, and sound articulation) and 11 statistical functions (such as maximum, minimum, range and mean). Table I shows the details of 44 acoustic features and 11 statistical functions.

In this paper, each dimensional feature for discrimination is newly weighted by employing a feature-weighting method based on Multiple Kernel Learning (MKL) in order to obtain the dimensional feature having information that is useful for classifying ASD. Then, the estimated dimensional feature for discrimination is classified by SVM.

The MKL [17] algorithm has been used to integrate multiple conventional kernel-based methods, such as SVMs, which rely only on a single kernel (see Fig. 2, upper panel) by assigning appropriate weights to those multiple component kernels. In a MKL framework, a combined kernel function is defined as a linear combination of the base kernel.

$$k(\mathbf{X}_i, \mathbf{X}_j) = \textstyle\sum_l \beta_l k_l(\mathbf{X}_i, \mathbf{X}_j) \quad (1)$$

Here k_l is the l-th base kernel computed from the i-th and j-th samples of the acoustic feature X_i and X_j, and the non-negative coefficient β_l represents the weight of the base kernel. The MKL approach for SVMs has been originally used to improve classifier performance by combining various classifiers with different kernels, each receiving the same feature vector (see Fig. 2, middle panel). Longworth [18] combines various dynamic kernels, including derivative kernel and parametric kernels, using MKL for speaker verification. In recent image recognition research, however, the MKL approach begun to be used for the purpose of feature vector selection or weighting.

For this purpose, the weight is independently trained

TABLE I
44 DIMENSIONAL ACOUSTIC FEATURES AND 11 STATISTICAL FUNCTIONS.

Acoustic features		dim.
mfcc	Mel–Frequency cepstral coefficients.	12
Δ mfcc	Delta Mel–Frequency cepstral coefficients.	12
lspFreq	Line spectral pair frequencies.	8
Δ lspFreq	Delta Line spectral pair frequencies.	8
pcm_zcr	Zero–crossing rate of time signal (frame–based).	1
voiceProb	Voicing probability computed from autocorrelation function.	1
F0	Fundamental frequency computed from the Cepstrum.	1
F0env	Envelope of the smoothed fundamental frequency contour.	1

Statistical functions	
Max	Maximum value of contour.
Min	Minimum value of contour.
Range	Max. – Min.
maxPos	Absolute position of maximum value.
minPos	Absolute position of minimum value.
Mean	Arithmetic mean of contour.
Linregc1	Slope (m) of a linear approximation of contour.
Linregc2	Offset (t) of a linear approximation of contour.
Stddev	Standard deviation of values in contour.
Skewness	Skewness (3rd order moment).
Kurtosis	Kurtosis (4th order moment).

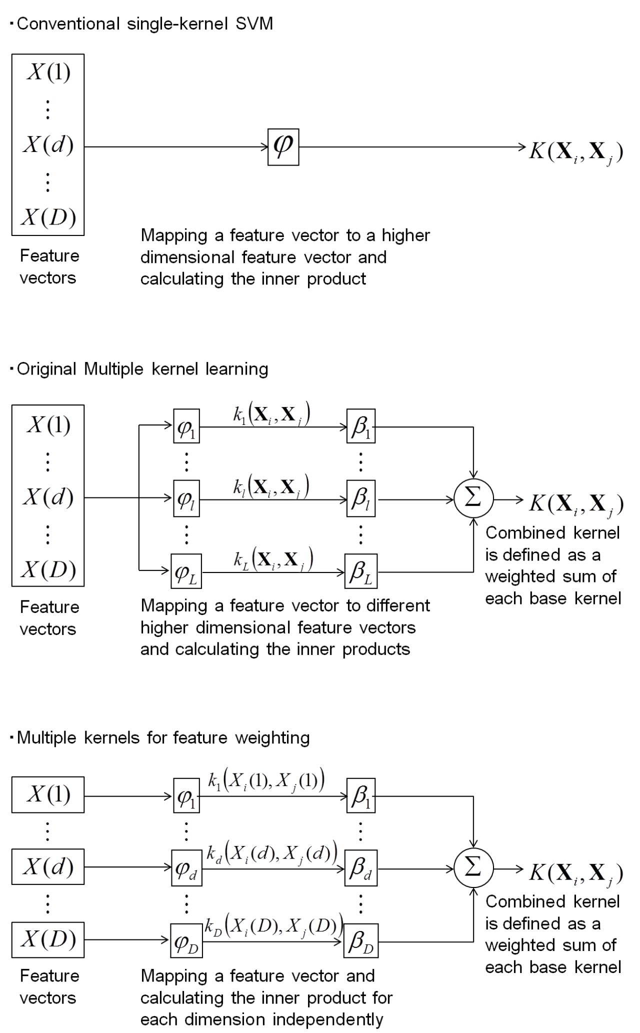

Fig. 2. A conventional SVM with single-kernel, original MKL for SVMs and a new weighting method based on MKL

for each base kernel receiving some different feature vectors [14]. In this paper, we propose a feature-weighting method for dimensional feature, where the weights of each dimensional feature are trained by MKL, defining the base kernels for each dimensional feature (see Fig. 2, lower panel).

$$k(\mathbf{X}_i,\mathbf{X}_j) = \textstyle\sum_d \beta_d k_d(X_i(d),X_j(d)) \tag{2}$$

The kernel weight β_d is trained in the SVM framework (i.e., maximum-margin based scheme). In the SVM framework, the MKL criterion is defined by the following objective function [17].

$$\max_{\alpha,\beta} \textstyle\sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i\alpha_j y_i y_j \sum_d \beta_d k_d(X_i(d),X_j(d))$$

$$s.t. \begin{cases} \sum_i y_i\alpha_i = 0, & 0 \le \alpha_i \le C \\ \sum_d \beta_d = 1, & \beta_d \ge 0 \end{cases} \tag{3}$$

Here α_i is the Lagrange coefficient, and $y_i = \{+1,-1\}$ denotes the class label of example X_i. C determines the trade-off between the margin and training data error. In (3), both α_i and β_d are optimized by a two-step iterative procedure. In the first step, β_d is fixed, and α_i is updated by a standard SVM solver. In the second step, α_i is fixed, and β_d is updated. In this paper, we use SVM^{light} [19] to obtain α_i, and optimize β_d by a projected-gradient scheme [17], [20]. In this way, the feature weights and the classification boundary are trained simultaneously.

III. EXPERIMENTS

A. Experiment Conditions

In our experiments, we recorded the speech data of children. The children ranged in age from kindergarten-

978-1-4799-2625-1/13/$31.00 ©2013 IEEE

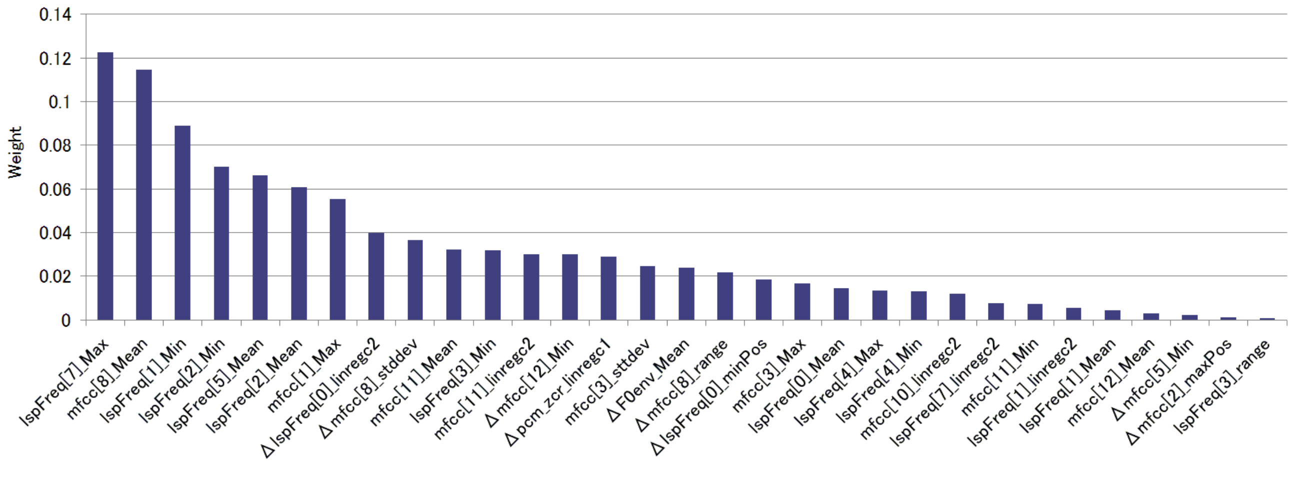

Fig. 3. Weight of features estimated by using MKL-SVM

ers to the fourth graders. This database consists 20 ASD children (1,560 words) and 21 typically developing children (1,508 words).

ASD participants were recruited from among the 4- to 9-year-old children who visited the Developmental Behavioral Pediatric Clinic of Kobe University Hospital between April and July 2010, after approval from the Medical Ethics Committee of Kobe University Graduate School of Medicine was received.

We divided this database into three and applied 3-fold cross-validation to our experiments. In 3-fold cross-validation, the original sample is randomly partitioned into 3 equal size subsamples. One subsample is used as the validation data for testing the model, and the other two subsamples are used as training data. The cross-validation process is then repeated three times (the folds), with each of the three subsamples used exactly once as the validation data. The average over the three results from the folds was then computed.

B. Experiment Results

Table II shows the classification accuracy for each method. The classification accuracy of MKL-SVM (80.2%) is 9.5% higher than that of SVM (70.7%). This result shows that MKL-SVM is effective in the classification of children with autism spectrum and typically developing children.

Fig. 3 and Table III show the weights of the best 31 features selected by MKL-SVM. Those results show that line spectral frequency, which is one of the formant features, is effective in the classification of children with autism spectrum and typically developing children.

TABLE II

CLASSIFICATION RESULTS

	SVM	SVM*	MKL-SVM
Number of dims.	484	31*	484→**31**
Feature selection	–	○	○
Weighting	–	–	○
Accuracy [%]	70.7	79.7*	**80.2**

In the Table II, SVM* means SVM using the best 31 features selected by MKL-SVM. The classification accuracy of SVM* (79.7%) is 9.0% higher than that of SVM using 484 features (70.7%) because of the feature selection. Also, the classification accurasy of MKL-SVM is 0.5% higher than that of SVM* due to the feature weighting (and selection).

IV. Conclusion

This paper described the result of a classification experiment carried out on children with autism spectrum and typically developing children from kindergarteners to the fourth graders. In our proposed method, a new acoustic feature weighting method was presented using a multiple-kernel learning algorithm for our discrimination task.

In our MKL approach, the weight is independently trained for each base kernel receiving some different feature vectors. We propose a feature-weighting method for dimensional feature, where the weights of each dimensional feature are trained by MKL, defining the base kernels for each dimensional feature.

TABLE III

Features selected by MKL-SVM

Order	Selected feature
1	Max. of the 7th ord. of lspFreq
2	Mean of the 8th ord. of mfcc
3	Min. of the 1st ord. of lspFreq
4	Min. of the 2nd ord. of lspFreq
5	Mean. of the 5th ord. of lspFreq
6	Mean. of the 2nd ord. of lspFreq
7	Max. of the 1st ord. of mfcc
8	Linear reg. coefficients of the 0th ord. of ΔlspFreq
9	Standard dev. of the 8th ord. of Δmfcc
10	Mean. of the 11th ord. mfcc
11	Min. of the 3rd ord. lspFreq
12	Linear reg. coefficients of the 11th ord. of mfcc
13	Min. of the 12th ord. Δmfcc
14	Linear reg. coefficients of Δpcm_zcr
15	Standard dev. of the 3th ord. of mfcc
16	Mean. of ΔF0env
17	Range of the 8th ord. of Δmfcc
18	Min. position of the 0th ord. of ΔlspFreq
19	Max. of the 3rd ord. of mfcc
20	Mean. of the 0th ord. of lspFreq
21	Max. of the 4th ord. of lspFreq
22	Min. of the 4th ord. of lspFreq
23	Linear reg. coefficients of the 10th ord. of mfcc
24	Linear reg. coefficients of the 7th ord. of lspFreq
25	Min. of the 11th ord. of mfcc
26	Linear reg. coefficients of the 1st ord. of lspFreq
27	Mean. of the 1st ord. of lspFreq
28	Mean. of the 12th ord. of mfcc
29	Min. position of the 5nd ord. of Δmfcc
30	Max. position of the 2nd ord. of Δmfcc
31	Range of the 3rd ord. lspfreq

The classification accuracy of MKL-SVM (80.2%) gives 9.5% higher than the accuracy of SVM (70.7%), and our results show the large weight acoustic features mainly in line spectral frequencies. Our results show that the line spectral frequency, which is one of the formant features, is effective in the classification of children with autism spectrum and typically developing children.

In the future, we will perform experiments that involve increasing the amount of speech data, considering other classification method (such as Boosting) and clarifying the influence line spectral frequencies have on the classification of children with autism spectrum and typically developing children.

References

[1] Ryo Aihara, Ryoichi Takashima, Tetsuya Takiguchi, and Yasuo Ariki, "Consonant enhancement for articulation disorders based on non-negative matrix factorization," in *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 4 pages*, 2012.

[2] Toshiya Yoshioka, Ryoichi Takashima, Tetsuya Takiguchi, and Yasuo Ariki, "Robust feature extraction to utterance fluctuations due to articulation disorders based on sparse expression," in *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 4 pages*, 2012.

[3] Lorna Wing, "Asperger's syndrome: a clinical account," *Psychological Medicine*, pp. 115–129, 1981.

[4] Lorna Wing, "Past and future of research on Asperger syndrome," *Asperger syndrome*, pp. 418–432, 2000.

[5] Elizabeth B. Caronna, Jeff M. Milunsky, and Helen Tager-Flusberg, "Autism spectrum disorders: clinical and research frontiers," *Archives of Disease in Childhood*, vol. 93, no. 6, pp. 518–523, 2008.

[6] Michael Rutter, "Incidence of autism spectrum disorders: Changes over time and their meaning," *Acta Paediatrica*, vol. 94, no. 1, pp. 2–15, 2005.

[7] Chris Plauché Johnson Scott M. Myers et al., "Management of children with autism spectrum disorders," *Pediatrics*, vol. 120, no. 5, pp. 1162–1182, 2007.

[8] Sally J. Rogers and Laurie A. Vismara, "Evidence-based comprehensive treatments for early autism," *Journal of Clinical Child & Adolescent Psychology*, vol. 37, no. 1, pp. 8–38, 2008.

[9] Ami Klin, David Pauls, Robert Schultz, and Fred Volkmar, "Three diagnostic approaches to Asperger syndrome: implications for research," *Journal of Autism and Developmental Disorders*, vol. 35, no. 2, pp. 221–234, 2005.

[10] Hideo Honda, Yasuo Shimizu, and Michael Rutter, "No effect of MMR withdrawal on the incidence of autism: a total population study," *Journal of Child Psychology and Psychiatry*, vol. 46, no. 6, pp. 572–579, 2005.

[11] Andy Bondy and Lori Frost, *A Picture's Worth: PECS and Other Visual Communication Strategies in Autism. Topics in Autism.*, ERIC, 2002.

[12] Shira Richman, *Raising a child with autism: a guide to applied behavior analysis for parents*, Jessica Kingsley Publishers, 2001.

[13] Robert Paul Liberman, William J DeRisi, and Kim T Mueser, *Social skills training for psychiatric patients.*, Pergamon Press, 1989.

[14] M. Varma and D. Ray, "Learning the discriminative power-invariance trade-off," in *Proc. ICCV07*, Rio de Janeiro, Brazil, October 2007, pp. 1–8.

[15] Florian Eyben, Martin Wollmer, and Björn Schuller, "OpenEAR-introducing the Munich open-source emotion and affect recognition toolkit," in *Proc. of the 4th International HUMAINE Association Conference on Affective Computing and Intelligent Interaction, 6 pages*, 2009.

[16] Björn Schuller, Stefan Steidl, and Anton Batliner, "The Interspeech 2009 emotion challenge," in *Proc. Interspeech*, 2009, pp. 312–315.

[17] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "More efficiency in multiple kernel learning," in *Proc. ICML*, Corvallis, Oregon, USA, June 2007, pp. 775–782.

[18] C. Longworth and M.J.F. Gales, "Multiple kernel learning for speaker verification," in *Proc. ICASSP08*, Las Vegas, Nevada, USA, March 2008, pp. 1581–1584.

[19] Bernhard Schölkopf and Christopher JC Burges, *Advances in kernel methods: support vector learning*, The MIT press, 1999.

[20] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.