A FUZZY FUSION ALGORITHM TO COMBINE MULTIPLE CLASSIFIERS

Bor-Chen Kuo [1] *Chih-Sheng Huang* [1] *Hsiang-Chuan Liu* [2] *Chih-Cheng Hung* [3]
kbc@mail.ntcu.edu.tw glanboy@hotmail.com lhc@asia.edu.tw chung@spsu.edu

[1] Graduate Institute of Educational Measurement and Statistics, National Taichung University, Taichung, Taiwan, R.O.C.
[2] Department of Bioinformatics Asia University, Taiwan, ROC
[3] Hendri Purnawan Southern Polytechnic State University Marietta, GA 30067 USA

1. INTRODUCTION

Combining multiple classifiers is a natural way to discover useful information and improve the performances of individual classifiers. It's based on the combination of the outputs of an ensemble of different classifiers. When interactions exist in combining multiple classifiers, fuzzy integral would be a valid method to fuse multiple classifiers. In this fuzzy fusion approach, the fuzzy measure plays an important role. Liu [1] proposed a novel fuzzy measure, *L*-measure, which is more sensitive than some common measures, like λ-measure, *P*-measure and ν-measure. In this paper, we would combine the multiple classifiers by Choquet integral wtih this *L*-measure.

2. FUZZY FUSION

2.1. A Choquet integral wtih fuzzy measure

A fuzzy measure μ on a finite set $\boldsymbol{X}$ is defined as a set function $\mu : 2^X \to [0, 1]$, if it satisfied following axioms:

(i) $\mu(\phi) = 0, \mu(\boldsymbol{X}) = 1$ (normality), (ii) $\forall \boldsymbol{A}, \boldsymbol{B} \in 2^X, \boldsymbol{A} \subseteq \boldsymbol{B} \Rightarrow \mu(\boldsymbol{A}) \le \mu(\boldsymbol{B})$ (monotonicity),

and $\mu_{\boldsymbol{L}}$ is a *L* measure [2] on a finite set $\boldsymbol{X}$, $|\boldsymbol{X}| = n$, $\mu_{\boldsymbol{L}}$ has to satisfy the following axioms:

(i) $\mu_{\boldsymbol{L}}(\phi) = 0$, $\mu_{\boldsymbol{L}}(\boldsymbol{X}) = 1$, $\boldsymbol{L} \in [0, \infty)$, (ii) $\forall \boldsymbol{A} \subset \boldsymbol{X}$, $n - |\boldsymbol{A}| + (|\boldsymbol{A}| - 1)\boldsymbol{L} > 0$,

$\mu_{\boldsymbol{L}}$ can express the formula as follow:

$$\mu_{\boldsymbol{L}}(\boldsymbol{A}) = \max_{\boldsymbol{x} \in \boldsymbol{A}}\left[\mu_{\boldsymbol{L}}(\{\boldsymbol{x}\})\right] + \frac{(|\boldsymbol{A}| - 1)L \sum_{x \in A} \mu_{\boldsymbol{L}}(\{\boldsymbol{x}\})}{\left[n - |\boldsymbol{A}| + (|\boldsymbol{A}| - 1)L\right] \sum_{\boldsymbol{x} \in \boldsymbol{X}} \mu_{\boldsymbol{L}}(\{\boldsymbol{x}\})} \left[1 - \max_{\boldsymbol{x} \in \boldsymbol{A}}\left[\mu_{\boldsymbol{L}}(\{\boldsymbol{x}\})\right]\right].$$

Let μ be a fuzzy measure on a finite set X, $h(\boldsymbol{x}_1) \le h(\boldsymbol{x}_2) \le \ldots \le h(\boldsymbol{x}_n)$ and $\boldsymbol{A_i} = \{\boldsymbol{x}_j, \boldsymbol{x}_{j+1}, \ldots, \boldsymbol{x}_n\}$. The Choquet integral of $\mathrm{h} : \mathrm{X} \to [0, 1]$ with respect to μ is denoted by

$$C_\mu(h) \overset{def}{=} \int h dg \overset{def}{=} h(x_1)\mu(A_1) + \sum_{\boldsymbol{i}=2}^{\boldsymbol{n}} (h(x_i) - h(x_{i-1}))\mu(\boldsymbol{A_i})$$

2.2. Fuzzy fusion algorithm with decision profile

Let $\boldsymbol{X} \in \boldsymbol{R^n}$ be a feature vector, $\boldsymbol{W} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w_C}\}$ be a label set of *C* classes. Each classifier D_i in the ensemble $\boldsymbol{D} = \{D_1, D_2, \ldots, D_R\}$ outputs *C* degrees of support. The degrees of support for a given input $\boldsymbol{x}$ are interpreted in different forms. In this study, the degrees of support for a given input $\boldsymbol{x}$ are defined as the posterior probabilities of the classes. Suppose $\mathrm{D_i} : R^n \to [0, 1]^C$, that is, all *C* degrees are assumed in the interval $[0, 1]$. The output of the *i*-th classifier is $D_i = [p_{i,1}(x), \cdots, p_{i,C}(x)]^T$, where $p_{i,j}(\boldsymbol{x})$ is the degree of support for a given sample $\boldsymbol{x}$ and assigned to the class j by the classifier D_i. In Figure 1, the posterior probabilities of $\boldsymbol{x}$ from R classifiers can be formed in a decision profile (DP) as the matrix.

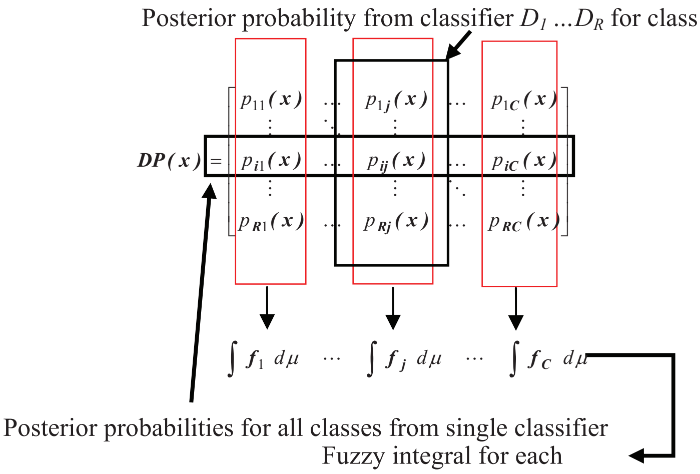

Figure 1. Decision profile(DP) matrix for the sample $\boldsymbol{x}$ and the fuzzy integral for the class

The decision rule of which class sample $\boldsymbol{x}$ should be assigned to is as following:

$$w(x) = \arg\max_{j=1:C} \{\int f_j(x)\, \mathrm{d}\mu_L(A_j)\},$$

where $\int f_j(x)\, \mathrm{d}\mu_L(A_j)$ represents the fused degree of support from all classifiers for class j by Choquet integral with L-measure. $f_j(x) = [p_{1j}(x), p_{2j}(x), \cdots, p_{Rj}(x)]^T$ and $A_j = [A_{1j}, A_{2j}, \cdots, A_{Rj}]^T$ are a set of posterior probability and a set of weights (fuzzy measures) for class j, respectively.

3. SOME EXPERIMENTAL RESULTS

In this study, there are two hyperspectral images are applied. One is Indian Pine as a mixed forest/agricultural site in Indiana and the other one is Washington, DC Mall as an urban site. Some results are shown in Table 1.

Table 1. Overall Accuracies for the case with 40 training samples per class in Indian Pine dataset

Classifier	Number of classifiers / Fusion method	10	20	30	40	50	60	70	80	90	100
ML	Majority Vote	0.828	0.852	0.826	0.776	0.726	0.692	0.643	0.618	0.592	0.581
	Choquet-Fuzzy	0.834	0.860	0.838	0.805	0.778	0.760	0.721	0.672	0.653	0.638

4. REFERENCES

[1] Kuncheva, L.I., *Combining Pattern Classifiers: Methods and Algorithms*, John Wiley & Sons, New Jersey,2004.

[2] H.C. Liu, Y.D. Jheng, W.C. Lin, and G.S. Chen, "A novel fuzzy measure and its Choquet Integral Regression Model," *Proceedings of International conference on Machine Learning and Cybernetics 2007*, Vol. 7,No. 3 pp. 1394-1398, 2007.

[3] A.A. Ahmed and D. Mohamed, "A New Technique for Combining Multiple Classifiers using The Dempster-Shafer Theory of Evidence," *JAIR*, vol. 17, pp. 333-361, Nov., 2002.