Weighted Mean Assignment of a Pair of Correspondences Using Optimisation Functions*

Carlos Francisco Moreno-García and Francesc Serratosa

Universitat Rovira I Virgili,
Departament d'Enginyeria Informàtica I Matemàtiques, Spain
carlosfrancisco.moreno@estudiants.urv.cat
francesc.serratosa@urv.cat

Abstract. Consensus strategies have been recently studied to help machine learning ensure better results. Likewise, optimisation in graph matching has been explored to accelerate and improve pattern recognition systems. In this paper, we present a fast and simple consensus method which, given two correspondences of sets generated by separate entities, enounces a final consensus correspondence. It is based on an optimisation method that minimises the cost of the correspondence while forcing it (to the most) to be a weighted mean. We tested our strategy comparing ourselves with the classical minimum cost matching system, using a palmprint database, with each palmprint is represented by an average of 1000 minutiae.

Keywords: Consensus strategy, Hamming Distance, Weighted Mean, Assignment Problem, Optimisation.

1 Introduction

When two subjects decide to solve the assignment problem, differences on the points' mapping may occur. These differences appear due to several factors. Between them, we could cite the following. One of the subjects gives more importance to some of the point attributes and the other subject believes other ones are more important. For instance, if the sets of points represent regions of segmented images, one subject may think the area is more important than the colour, and the other one can think it is the opposite. If the assignment problem is solved by an artificial system, the fact of "believing" the area is more important than the colour is gauged by some weights. Another factor could be that the assignment problem is computed in a suboptimal algorithm, and different non-exact assignments can appear. In these scenarios, a system can intervene as a third party to decide the final assignment as a consensus of both assignments since some discrepancies will appear, especially as the number of involved points increase.

This paper presents a method to find the consensus assignment between two sets given two different assignments between those sets. We model the consensus

* This research is supported by the Spanish CICYT project DPI2013-42458-P and Consejo Nacional de Ciencia y Tecnologías (CONACyT Mexico).

P. Fränti et al. (Eds.): S+SSPR 2014, LNCS 8621, pp. 301–311, 2014.

assignment as the weighted mean assignment. This method is inspired in the one presented in [1] that obtains a clustering consensus given a set of clusterings. Other methods to perform this task are [2], where a final cluster is obtained based on a similarity graph and [3], where the least square algorithm is used. Our method, and also the one in [1], does not restrict the consensus assignment to be a strict mean but a weighted mean. This occurs because these methods aim to find an assignment (or clustering in [1]) that it is as closer as possible to both assignments (clusterings in [1]), but have to minimise the assignment cost (or the clustering cost in [1]). These methods are closely related to the unsupervised machine learning methods [4].

The drawback of this approach resides in the large number of possible solutions. One of the most well-known and practical options to reduce the complexity of a combinatorial calculation is combinatorial optimisation. The concept of optimisation is related to the selection of the "best" configuration or set of parameters to achieve a certain goal [5]. Functions involved in an optimisation problem can be either conformed by continuous values or discrete values, often called "Combinatorial Scenarios". These second scenarios have been largely studied and applied for graph matching problems, particularly in the case of the Hungarian Algorithm [6]. This method converts a combinatorial problem into an assignment problem, which will eventually derive in an optimal configuration for a cost-based labelling. Many recent researches have used graph theory and optimisation to solve diverse problems. Examples can be found in [7], where a graph representation and an optimisation method helped to design a gas drainage system for a coal mine. On [8], a research group used graph representation of newspaper articles to optimize the arrangement of each article within the page. Energy reduction in machinery [9] and most recently, biomedical compounds represented as labelled graphs [10] have been classified by using optimisation methods.

2 Basic Definitions

Given a set of elements $\mathrm{G}^1 = \{\mathrm{g}_1^1, \mathrm{g}_2^1, \ldots, \mathrm{g}_\mathrm{u}^1\}$, where the elements posses $\mathrm{g}_\mathrm{i}^1 = (\mathrm{m}_\mathrm{i}^1, \mathrm{a}_\mathrm{i}^1)$, being $\mathrm{m}_\mathrm{i}^1 \in \Sigma$ (where Σ is a unique number of the elements) and $\mathrm{a}_\mathrm{i}^1 \in \mathrm{T}$ (where T is the domain of the attribute of the elements), a labelling function f can be established between G^1 and another set of elements with similar characteristics G^2. This labelling function f is understood as a bijective function that proposes a match $f: \Sigma \rightarrow \Sigma$ from one element of G^1 to one element of G^2, where G^1 and G^2 have similar cardinality u.

We define the cost of a labelling $Cost(G^1, G^2, f)$ as the addition of individual element costs in a similar way as in the Graph Edit Distance [11],

$$Cost(f) = \sum_{i=1}^{u} c\left(a_i^1, a_j^2\right) \quad being\ j\ such\ that\ f(m_i^1) = m_j^2 \tag{1}$$

where c is defined as a distance function over the domain of attributes T and is application dependent [11].

The distance between sets $d_S(\cdot)$, which also delivers the minimum cost of all the labellings, is a function defined as

$$d_S(G^1, G^2) = min\{Cost(G^1, G^2, f)\}\ \forall\ f: \Sigma^1 \rightarrow\ \Sigma^2 \tag{2}$$

The labelling that obtains this distance is known as the optimal labelling f^*, and it is defined as

$$f^* = argmin_{\forall f: \Sigma^1 \rightarrow \Sigma^2}\{Cost(G^1, G^2, f)\} \tag{3}$$

We convert this linear minimisation problem into an assignment problem [6], for which any labelling f is related with a combination. With the calculation of a cost matrix $\boldsymbol{C}[i,j] = Cost(\mathrm{a}_\mathrm{i}^2, \mathrm{a}_\mathrm{j}^2)$, we can convert equation 3 into

$$f^* = argmin_{\forall f: \Sigma^1 \rightarrow \Sigma^2}\{\mathbf{C}_f\} \tag{4}$$

where $\mathbf{C}_f$ is the cost of the combination$f$ (or labelling in the set domain) applied to matrix $\boldsymbol{C}$. That is

$$\mathbf{C}_f = \sum_{i=1}^{\mathrm{u}} \boldsymbol{C}[i,k]\ where\ f(m_i^1) = m_k^2 \tag{5}$$

Assume f^a and f^b are two labelling functions between sets $\mathrm{G}^1 = \{\mathrm{g}_1^1, \mathrm{g}_2^1, \ldots, \mathrm{g}_\mathrm{u}^1\}$ and $\mathrm{G}^2 = \{\mathrm{g}_1^2, \mathrm{g}_2^2, \ldots, \mathrm{g}_\mathrm{u}^2\}$. We then define the Hamming Distance $\mathrm{d}_\mathrm{H}(\cdot)$ between the labellings $\mathrm{f}^\mathrm{a}$and f^b as

$$d_H(f^a, f^b) = \sum_{i=1}^{u}\left(1 - \partial\left(m_x^2, m_y^2\right)\right) \tag{6}$$

being x and y such that $\mathrm{f}^\mathrm{a}(\mathrm{m}_\mathrm{i}^1) = \mathrm{m}_\mathrm{x}^2$ and $\mathrm{f}^\mathrm{b}(\mathrm{m}_\mathrm{i}^1) = \mathrm{m}_\mathrm{y}^2$. Function ∂ is the well-known function known as the Kronecker Delta.

$$\partial(\mathrm{a},\mathrm{b}) = \begin{cases} 0 & \text{if } \mathrm{a} \neq \mathrm{b} \\ 1 & \text{if } \mathrm{a} = \mathrm{b} \end{cases} \tag{7}$$

In its more general form, the mean of two elements e^a and e^b has been defined as an element $\bar{\mathrm{e}}$ such that $\mathrm{d}(\mathrm{e}^\mathrm{a}, \bar{\mathrm{e}}) = \mathrm{d}\left(\mathrm{e}^\mathrm{b}, \bar{\mathrm{e}}\right)$ and $\mathrm{d}\left(\mathrm{e}^\mathrm{a}, \mathrm{e}^\mathrm{b}\right) = \mathrm{d}(\mathrm{e}^\mathrm{a}, \bar{\mathrm{e}}) + \mathrm{d}\left(\bar{\mathrm{e}}, \mathrm{e}^\mathrm{b}\right)$, being d any distance measure defined on the domain of these elements. Moreover, the weighted mean is sometimes used to gauge the importance or the contribution of the involved elements. In this case, the most general definition is $\mathrm{d}(\mathrm{e}^\mathrm{a}, \bar{\mathrm{e}}) = \alpha$ and $\mathrm{d}\left(\mathrm{e}^\mathrm{a}, \mathrm{e}^\mathrm{b}\right) = \alpha + \mathrm{d}\left(\bar{\mathrm{e}}, \mathrm{e}^\mathrm{b}\right)$ where α is a constant that controls the contribution of the elements and holds $0 \leq \alpha \leq \mathrm{d}\left(\mathrm{e}^\mathrm{a}, \mathrm{e}^\mathrm{b}\right)$. Finally, if we do not need to introduce the weighting factor α in our model, we have that any element $\bar{\mathrm{e}}$ is a weighted mean of two elements e^a and e^b if it holds that $\mathrm{d}\left(\mathrm{e}^\mathrm{a}, \mathrm{e}^\mathrm{b}\right) = \mathrm{d}(\mathrm{e}^\mathrm{a}, \bar{\mathrm{e}}) + \mathrm{d}\left(\bar{\mathrm{e}}, \mathrm{e}^\mathrm{b}\right)$. Note that elements e^a and e^b hold this condition so, they are also weighted means of themselves.

The most appropriate form to model a consensus scenario given two different options is the one done by [4], which is defined through the weighted mean of these two options. The aim of [4] is to find the consensus clustering of a set of elements given two different clustering proposals applied to this set of elements. If we want to translate this model to our problem, we should find the weighted mean labelling $\bar{\mathrm{f}}$ given two different labelling f^{a} and f^{b}. As commented in the previous paragraph, if we want $\bar{\mathrm{f}}$ to be defined as a weighted mean labelling of f^{a} and f^{b} the following restriction has to hold,

$$d_H(f^a, f^b) = d_H(f^a, \bar{f}) + d_H(\bar{f}, f^b) \tag{8}$$

Several labellings $\bar{f}$ hold this condition, and amongst them are f^a and f^b. To select one, an option would be a brute force method that obtains all possible combinations and selects the best one from the application point of view. Another option is a standard minimisation approach, to reduce the computational time.

Standard minimisation approaches aim to find an optimal element e^* that globally minimises a specific function. Usually this function is composed of an empirical risk $\nabla(\mathrm{e})$ plus a regularization term $\Omega(\mathrm{e})$ weighted by a parameter λ [5]. The empirical risk is the function to be minimised per se and the regularisation term is a mathematical mechanism to impose some restrictions. Parameter λ weights how much these restrictions have to be imposed.

$$e^* = \mathrm{argmin}_{\forall e} \{\nabla(e) + \lambda \cdot \Omega(e)\} \tag{9}$$

The aim of this paper is to present a method to find an approximation of the weighted mean labelling given two labellings. Therefore, we want to find $\bar{f}^*$ such that the following equation holds,

$$\bar{f}^* = \mathrm{argmin}_{\forall f: \Sigma^1 \to \Sigma^2} \{\lambda_C \cdot \nabla(f) + \lambda_H \cdot \Omega(f)\} \tag{10}$$

On the next section, we explain functions $\nabla(f)$ and $\Omega(f)$. Although it is not strictly necessary, we present this general equation with parameters λ_C and λ_H, instead of only one parameter λ as in equation 9, to simplify some explanations and examples.

3 Method

Our method defines the optimal labelling $\bar{f}^*$ through equation 10 in which the Loss function and the Regularisation term are

$$\nabla(f) = Cost(G^1, G^2, f) \text{ and } \Omega(f) = d_H(f^a, f) + d_H(f, f^b) - d_H(f^a, f^b) \tag{11}$$

That is, we want to minimise the labelling cost (equation 1) of the obtained labelling but restricted to be a weighted mean (equation 8). The degree of restriction depends on weights λ_C and λ_H. Note that by definition of a distance, $d_H(f^a, f) + d_H(f, f^b) - d_H(f^a, f^b) \geq 0$.

The aim of our method is to decide the labelling closer to both human's labellings. Therefore it seems logical that our strategy only seeks for the partial labelling where both of the specialists disagree. The other partial labelling, which is the one that both specialist has decided the same point mapping, is directly assigned as the mappings of these specialists. For this reason we split labellings f^a and f^b in two disjoint partial labellings such that $f^a = f'^a \cup f''^a$ and $f^b = f'^b \cup f''^b$, where f'^a and f'^b are the partial labellings where $f^a(m_i^1) = f^b(m_i^1)$, and f''^a and f''^b are the other partial ones where $f^a(m_i^1) \neq f^b(m_i^1)$. This also means that the cost of both labellings is $\mathrm{Cost}(G^1, G^2, f^a) = \mathrm{Cost}(G^1, G^2, f'^a) + \mathrm{Cost}(G^1, G^2, f''^a)$ and $\mathrm{Cost}(G^1, G^2, f^b) = \mathrm{Cost}(G^1, G^2, f'^b) + \mathrm{Cost}(G^1, G^2, f''^b)$. We define $\Sigma = \Sigma' \cup \Sigma''$. The set of nodes Σ' in G^1 is composed of the nodes such that $f^a(m_i^1) = f^b(m_i^1)$ and the set of nodes Σ'' in G^1 is composed of the nodes such that $f^a(m_i^1) \neq f^b(m_i^1)$.

Thus, we define the weighted mean labelling $\bar{f}^*$ we want to obtain as a union of two partial labellings, $\bar{f}^* = \bar{f}'^* \cup \bar{f}''^*$ where $\bar{f}'^* = f'^a$ (which is the same than $\bar{f}'^* = f'^b$) and $\bar{f}''^*$ is the one defined in the following equation,

$$\bar{f}''^*_{\lambda_C,\lambda_H} = \mathrm{argmin}_{\forall f'':\, \Sigma''^1 \to \Sigma''^2} \{ \lambda_C \cdot \mathrm{Cost}(G^1, G^2, f'') + \lambda_H \cdot \Big(d_H\big(f''^a, f''\big) + d_H\big(f'', f''^b\big) - d_H\big(f''^a, f''^b\big) \Big) \} \quad (12)$$

To solve equation 12, we translate the linear minimisation problem to an assignment problem [5] as we have shown in equation 4, but instead of the cost matrix $\boldsymbol{C}$, our method minimises matrix H_{λ_C,λ_H} defined as follows,

$$H_{\lambda_C,\lambda_H} = \lambda_C \cdot C'' + \lambda_H \cdot [\mathbf{1} - F''^{a,b}] \quad (13)$$

where $C''[i,j] = c(m_i^1, m_j^2)$ being $m_i^1 \in \Sigma''^1$ and $m_j^2 \in \Sigma''^2$. Besides, $F''^{a,b} = F''^a + F''^b$, where F''^a and F''^b are the labelling matrices corresponding to f''^a and f''^b, respectively. Aditionally, $\mathbf{1}$ is a matrix of all ones. Note that the number of rows and columns of matrices C'', F''^a and F''^b is lower or equal than C. As more similar node mappings of f^a and f^b are, the smaller the number nodes in Σ''^1 and Σ''^2 is, and so, the dimensions of C'', F''^a and F''^b. This fact affects directly on the computational cost. In a practical application, if both specialists are good enough, they discern in few node mappings and therefore the computational time of finding the agreement labelling is very low. Considering equation 13, we obtain the following expression,

$$\bar{f}''^*_{\lambda_C,\lambda_H} = \mathrm{argmin}_{\forall f''} \left\{ \left(H_{\lambda_C,\lambda_H} \right)_{f''} \right\} \quad (14)$$

Several algorithms can be used to minimise equation 14, for instance the Hungarian algorithm [5]. Finally, the cost of the obtained weighted mean becomes,

$$C_{\bar{f}^*_{\lambda_C,\lambda_H}} = C''_{\bar{f}''^*_{\lambda_C,\lambda_H}} + Cost\big(G^1, G^2, f'^a\big) \quad (15)$$

On section 3.1 we demonstrate equations 12 and 14 minimise at the same approximation of the weighted mean labelling $\bar{f}''^*_{\lambda_C,\lambda_H}$ for all weights λ_C and λ_H and

pair of graphs G^1 and G^2. Then on section 3.2 we demonstrate the cases in which the obtained labelling is an exact weighted mean labelling and not an approximated weighted mean labelling.

3.1 Reasoning about Optimality

If we want to use equation 14 to solve our problem instead of equation 12, we must now demonstrate that functional $\left\{\lambda_C \cdot \text{Cost}(G^1, G^2, f'') + \lambda_H \cdot \left(d_H(f''^a, f'') + d_H(f'', f''^b) - d_H(f''^a, f''^b)\right)\right\}$ extracted from equation 12 minimises the same partial labelling than $\left\{\left[\lambda_C \cdot C'' + \lambda_H \cdot [\mathbf{1} - F''^{a,b}]\right]_{f''}\right\}$ extracted from equation 14. Notice that, by definition, $Cost(G^1, G^2, f'') = C''_{f''}$ and for this reason, we have to demonstrate the following equation

$$[\mathbf{1} - F''^{a,b}]_{f''} = d_H(f''^a, f'') + d_H(f'', f''^b) - d_H(f''^a, f''^b); \forall f'': \Sigma''^1 \rightarrow \Sigma''^2 \quad (16)$$

If equation 16 holds, then we can confirm that it is valid to use equation 14 to solve our problem. Suppose the cardinality of $\Sigma''^1$and Σ''^2 is n. Therefore, by definition of these sets, $d_H(f''^a, f''^b) = n$. Given the involved labellings f'', f''^a and f''^b, we can define the three following natural numbers n_p, n_q and n_t:

1) n_p: number of nodes in Σ''^1that hold $f''(m_i^1) \neq f''^a(m_i^1)$ and $f''(m_i^1) \neq f''^b(m_i^1)$.
2) n_q: number of nodes in Σ''^1that hold $f''(m_i^1) = f''^a(m_i^1)$ and $f''(m_i^1) \neq f''^b(m_i^1)$.
3) n_t: number of nodes in Σ''^1that hold $f''(m_i^1) \neq f''^a(m_i^1)$ and $f''(m_i^1) = f''^b(m_i^1)$.

Again, by definition of these sets, there is not any m_i^1 such that $f''(m_i^1) = f''^a(m_i^1)$ and $f''(m_i^1) = f''^b(m_i^1)$. Therefore, $n = n_p + n_q + n_t$. By simplicity of notation, we order the nodes in Σ''^1 such that m_1^1 to $m_{n_p}^1$ hold the first condition, $m_{n_p+1}^1$ to $m_{n_p+n_q}^1$ hold the second condition and $m_{n_p+n_q+1}^1$ to m_n^1 hold the third condition.

To demonstrate that equation 16 holds, we first demonstrate that $[\mathbf{1} - F''^{a,b}]_{f''} = n_p$ and we second demonstrate that $d_H(f''^a, f'') + d_H(f'', f''^b) - d_H(f''^a, f''^b) = n_p$.

1) Demonstration of $[\mathbf{1} - F''^{a,b}]_{f''} = n_p$: Suppose that $f''(m_i^1) = m_k^2$ then $[\mathbf{1} - F''^{a,b}]_{f''} = \sum_{i=1}^{n}(\mathbf{1} - F''^{a,b})[i,k] = \sum_{i=1}^{n_p} 1 + \sum_{i=n_p+1}^{n} 0 = n_p$.
2) Demonstration of $d_H(f''^a, f'') + d_H(f'', f''^b) - d_H(f''^a, f''^b) = n_p$: $d_H(f''^a, f'') + d_H(f'', f''^b) - n = \sum_{i=1}^{n_p}\left(2 - \partial\left(f''^a(m_i^1), f''(m_i^1)\right) - \partial\left(f''(m_i^1), f''^b(m_i^1)\right)\right) - n = \left(\sum_{i=1}^{n_p} 0 + \sum_{i=n_p+1}^{n_p+n_q} 1 + \sum_{i=n_p+n_q+1}^{n} 1\right) - n = n_q + n_t - n = n_p$.

■

3.2 Exact Weighted Mean Labelling

In some cases, it is interesting to know if we have obtained an exact weighted mean labelling or an approximated one. First of all, we have to realise that whether both f' and f'' are exact partial weighted mean labellings, then the union $f = f' \cup f''$ is an exact weighted mean labelling, assuming that that $\mathrm{f}' \cap \mathrm{f}'' = 0$. It is clear that if equation 8 holds for both partial labellings then it holds for the complete one. Moreover, by definition of our partial labelling f', it is always defined as weighted mean labelling. Therefore, we conclude that the obtained labelling $\bar{\mathrm{f}}^*$is an exact weighted mean labelling if $\bar{\mathrm{f}}''^*$ is also an exact weighted mean labelling. These cases are the ones that that $\bar{\mathrm{f}}''^*$ holds equation 8. Due to we have demonstrated that $d_H(f''^a, f'') + d_H(f'', f''^b) - d_H(f''^a, f''^b) = \mathrm{n_p}$, then $\mathrm{n_p}$ has to be 0. By definition of $\mathrm{n_p}$, these labelling are the ones that $\bar{f}''^*(m_i^1) = f''^a(m_i^1)$ or $\bar{f}''^*(m_i^1) = f''^b(m_i^1)$. Therefore, we conclude the following expression,

$$\bar{\mathrm{f}}^*_{\lambda_C,\lambda_H} \text{is a weighted mean labelling if:}$$
$$\bar{\mathrm{f}}''^*_{\lambda_C,\lambda_H}(\mathrm{m_i^1}) = \mathrm{f}''^{\mathrm{a}}(\mathrm{m_i^1}) \text{ or } \bar{\mathrm{f}}''^*_{\lambda_C,\lambda_H}(\mathrm{m_i^1}) = \mathrm{f}''^{\mathrm{b}}(\mathrm{m_i^1}); \forall \mathrm{m_i^1} \in \Sigma''^1 \qquad (17)$$

The cost of testing if the labelling obtained is a weighted mean is linear on the number of discordances between labellings f^a and f^b.

Note that if $\lambda_C = 0$ and $\lambda_H > 0$ and we use the Hungarian method [6] to solve equation 14, then the method always obtains an exact weighted mean labelling. This is because equation 16 has been demonstrated, and also because the optimality of the Hungarian method has been demonstrated in equation 17.

4 Experimentation

We used images contained in the Tsinghua 500 PPI Palmprint Database [12]. It is a public high-resolution palmprint database composed of 500 palmprint images of a resolution of 2040 x2040 pixels. From each person, 8 palmprints are enrolled. We extracted the minutiae set of the 8 palmprints of the first 10 subjects of the database using the algorithm presented in [13], [14] and [15] and we obtained an average of 1000 minutiae per palmprint. The attributes of minutiae are their position and angle, which means $\mathrm{a_i^s} = \{\theta_\mathrm{i}^\mathrm{s}, (\mathrm{x}, \mathrm{y})_\mathrm{i}^\mathrm{s}\}$. Therefore, our core database is composed by 80 sets of minutiae classified in 10 subjects. Nevertheless, we wish to have a database of several registers and each register composed of four elements: 2 minutiae sets G^1 and G^2 extracted from the same palm and two different labellings f^a and f^b between these minutiae sets. To do so, we matched each of the 8 minutiae sets of the same subject obtaining 64 correspondences per subject. Therefore, we defined an initial database of 8x8x10=640 registers composed of 2 minutiae sets G^1 and G^2 extracted from the same subject and a correspondence $\check{\mathrm{f}}$ between them. Correspondences were

computed through the Hungarian method [16] and a greedy method to select the matches from the resulting matrix. The distance between minutiae has been defined to be

$$c\left(a_i^1, a_j^2\right) = 0.5 \cdot ad\left(\theta_i^1, \theta_j^2\right) + 0.5 \cdot dd\left((x,y)_i^1, (x,y)_j^2\right) \quad (18)$$

being *ad* the angular distance and *dd* the Euclidean distance.

Given each of the 640 registers, we need to generate labellings f^a and f^b from the initial labellings $\check{f}$. Nevertheless, to perform our experiments, we need to control the distance between these labellings. Therefore, we introduce parameter α that decides the Hamming distance between them, $d_H(f^a, f^b) = 2\alpha$. $f^a$and f^b are randomly generated such that $d_H\left(f^a, \check{f}\right) = \alpha$ and $d_H\left(f^b, \check{f}\right) = \alpha$. As we will see later, parameter α is the horizontal axis of the figures presented in this section and for each $\alpha \in \{10,11,\dots,212\}$we have a dataset of 640 registers, so, the values presented in these figures are the average of 640 times we computed $\bar{f}^*_{\lambda_C,\lambda_H}$.

The aim of our method is to find the consensus labelling with the minimum cost and close to both labellings. For this reason, we have performed several tests using different configurations of λ_C and λ_H and parameter α. The aim of these tests is threefold. First, we want to know the cost of the obtained labellings $\bar{f}^*_{\lambda_C,\lambda_H}$, that is $Cost\left(\bar{f}^*_{\lambda_C,\lambda_H}\right)$. Second, we want to analyse if the obtained labellings appear to be "in the middle" of both labellings. In this case, we propose the following measure,

$$\text{Middle}\left(f^a, f^b, \bar{f}^*_{\lambda_C,\lambda_H}\right) = \frac{\left|d_H\left(f^a, \bar{f}^*_{\lambda_C,\lambda_H}\right) - d_H\left(f^b, \bar{f}^*_{\lambda_C,\lambda_H}\right)\right|}{d_H(f^a, f^b)} \quad (19)$$

and third, we want to check if the obtained labellings are really weighted mean labellings.

For this experimentation we chose three different configurations for λ_H and λ_C. First, when $\lambda_H = 0$ and $\lambda_C = 1$ the labellings are not being considered, thus basing the decision only on the minimum cost. Therefore, this configuration will reproduce a classical minimum-cost method (red in Figures 1 to 3). Second, when $\lambda_H = 1$ and $\lambda_C = 1$ there is a contribution both of the cost and the labelling. Therefore, this approach would represent our method (green in Figures 1 to 3). Finally, when $\lambda_H = 1$ and $\lambda_C = 0$ only the labellings are being considered but no cost is used. Therefore this approach would be considered a pure consensus of the correspondences done by $f^a$and f^b(violet in Figures 1 to 3).

Figure 1 shows the cost of $\bar{f}^*_{\lambda_C,\lambda_H}$as the number of mistakes increases. Notice that the y-axis represents the cost (equation 1, where the cost between nodes is equation 18), being an application-dependent metric. However, it is clearly noticeable that the classical method (red) performs just as good as our method in terms of minimising the cost. It must be pointed out that a minimum cost not necessarily translates in a better result

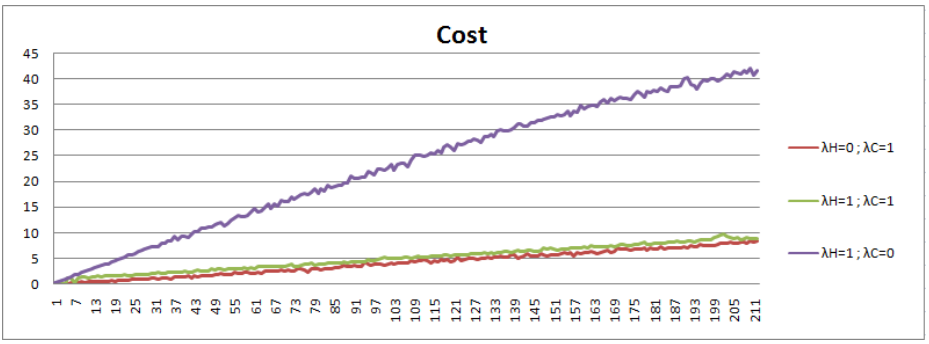

Fig. 1. Comparison for the cost of $\bar{f}^*_{\lambda_C,\lambda_H}$ for the three configurations

Figure 2 shows the Middle measure (equation 19) which measures how far in terms of Hamming Distance is the consensus labelling with respect of the median of f^a and f^b. We can see once again that the classical method (red) performs slightly worse than our approach, however as mistakes increase, the distance "Middle" decreases and eventually stabilizes for every configuration.

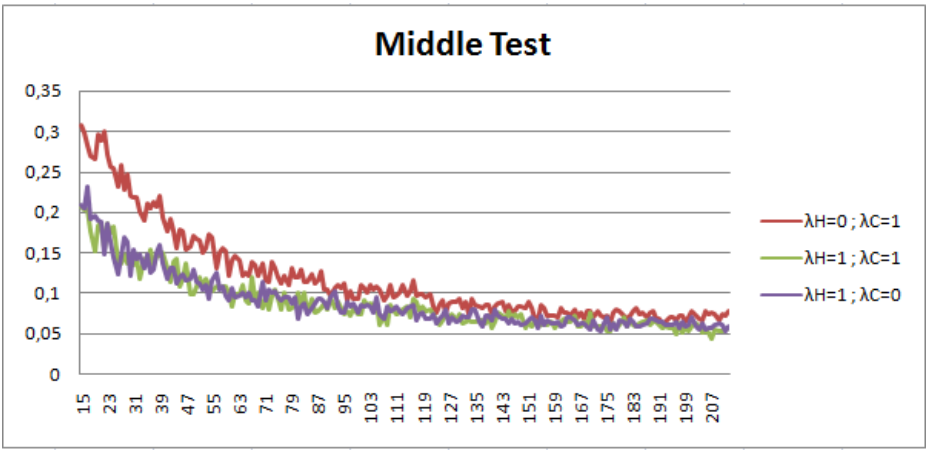

Fig. 2. Comparison for the position of $\bar{f}^*_{\lambda_C,\lambda_H}$ for the three configurations

Figure 3 shows the percentage of experiments in which $\bar{f}^*_{\lambda_C,\lambda_H}$ is really a weighted mean of f^a and f^b. We can notice that the classical method (red) does not always give weighted means starting from 25 mistakes (which means that it stops delivering results that are consensus). As we deducted from equation 17, the labelling-only configuration (violet) will always result in weighted means, whereas the approach that equally considers both terms (green) will slightly decrease in successful weighted mean results as α increase.

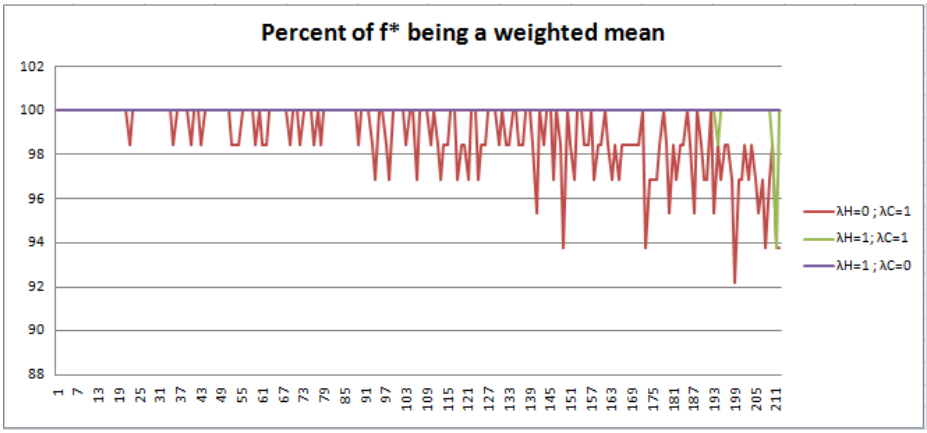

Fig. 3. Comparison for the percent of $\bar{f}^*_{\lambda_C,\lambda_H}$ being a weighted mean for the three configurations

5 Conclusions and Further Work

We present a fast and efficient method to perform a consensus decision based on a regularization term consisting in two labellings developed by two separate entities, and a loss function consisting on a cost relation which is application dependant. We also demonstrate that an optimisation process can be applied to reduce the computational cost of calculating the multiple possibilities and that different configurations on the loss function and the regularisation term can be produced to obtain different results. As a further work, we would like to continue studying the effects of consensus techniques in with multiple inputs and multiple metrics [17].

References

1. Franek, F., Jiang, X., He, C.: Weighted Mean of a Pair of Clusterings. Pattern Analysis Applications. Springer (2012)
2. Mimaroglu, S., Erdil, E.: Combining multiple clusterings using similarity graphs. Pattern Recognition 44, 694–703 (2010)
3. Murino, L., Angelini, C., De Feis, I., Raiconi, G., Tagliaferri, R.: Beyond classical consensus clustering: The least squares approach to multiple solutions. Pattern Recognition Letters 32, 1604–1612 (2011)
4. Ghahramani, Z.: Unsupervised Learning. Advanced Lectures on Machine Learning, pp. 72–112. Springer (2004)
5. Papadimitriou, C., Steiglitz, K.: Combinatorial Optimization: Algorithms and Complexity. Dover Publications (July 1998)
6. Kuhn, H.W.: The Hungarian method for the assignment problem Export. Naval Research Logistics Quarterly 2(1-2), 83–97 (1955)
7. Qing, Y., Haizhen, W., Zhenzhen, J., Yan, P.: Graph Theory and its application in optimization of gas drainage system in coal mine. Procedia Engineering 45, 339–344 (2012)

8. Gao, L., Wang, Y., Tang, Z., Lin, X.: Newspaper article reconstruction using ant colony optimization and bipartite graph. Applied Soft Computing 13, 3033–3046 (2013)
9. Eberspächer, P., Verl, A.: Realizing energy reduction of machine tools through a control-integrated consumption graph-based optimization method. Procedia CIRP 7, 640–645 (2013)
10. Livi, L., Rizzi, A., Sadeghian, A.: Optimized dissimilarity space embedding for labeled graphs
11. Solé, A., Serratosa, F., Sanfeliu, A.: On the Graph Edit Distance cost: Properties and Applications. International Journal of Pattern Recognition and Artificial Intelligence. IJPRAI 26(5) (2012)
12. Jain, A.K., Feng, J.: Latent Palmprint Matching. IEEE Trans. on PAMI (2009)
13. Funada, J., et al.: Feature Extraction Method for Palmprint Considering Elimination of Creases. In: Proc. 14th Int. Conf. Pattern Recognition, pp. 1849–1854 (1998)
14. Dai, J., Zhou, J.: Multifeature-Based High-Resolution Palmprint Recognition. IEEE Trans. Pattern Analysis and Machine Intelligence 33(5), 945–957 (2011)
15. Dai, J., Feng, J., Zhou, J.: Robust and Efficient Ridge Based Palmprint Matching. IEEE Transactions on Pattern Analysis and Matching Intelligence 34(8) (2012)
16. Kuhn, H.W.: The Hungarian Method for the assignment problem. Naval Research Logistics Quarterly 2, 83–97 (1955)
17. Franek, L., Jiang, X.: Ensemble clustering by means of clustering embedding in vector space. Pattern Recognition 47(2), 833–842 (2014)