

On the combination of two visual cognition systems using combinatorial fusion

Amy Batallones · Kilby Sanchez · Brian Mott ·
Cameron Coffran · D. Frank Hsu

Received: 17 May 2014 / Accepted: 8 January 2015 / Published online: 3 February 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract When combining decisions made by two separate visual cognition systems, statistical means such as simple average (M_1) and weighted average (M_2 and M_3), incorporating the confidence level of each of these systems have been used. Although combination using these means can improve each of the individual systems, it is not known when and why this can happen. By extending a visual cognition system to become a scoring system based on each of the statistical means M_1 , M_2 , and M_3 respectively, the problem of combining visual cognition systems is transformed to the problem of combining multiple scoring systems. In this paper, we examine the combined results in terms of performance and diversity using combinatorial fusion, and study the issue of when and why a combined system can be better than individual systems. A data set from an experiment with twelve trials is analyzed. The findings demonstrated that combination of two visual cognition systems, based on weighted means M_2 or M_3 , can improve each of the individual systems only when both of

them have relatively good performance and they are diverse.

Keywords Combinatorial fusion analysis (CFA) · Decision-making · Visual cognition · Rank-score characteristics (RSC) function · Cognitive diversity

1 Introduction

Many decisions that humans have to make are partially, or even wholly, based on visual input. The split second nature of such decisions may make the process seem simple. However, there are many factors that are considered and combined during this short time frame. On a neurological level, there has been growing interest in understanding the factors that are combined within the visual aspect alone [1, 2], as well as how visual information is joined with information from other senses [3–7]. Combination of multiple visual decisions has also been explored [5, 8, 9].

Prior research into how pairs of people can interactively make decisions based on visual perception has been conducted by several researchers including Bahrami et al. [8], Ernst [5], and Kepecs et al. [9]. In Bahrami's work, four predictive models are used on experiments of varying degrees of noise, feedback, and communication: coin-flip (CF), behavioral feedback (BF), weighted confidence sharing (WCS), and direct signal sharing (DSS). Bahrami concludes that the WCS model is the only one that can fit over the empirical data. His findings indicate that the accuracy of the decision-making is aided by communication between the pairs and can greatly improve the overall performance of the pair.

Marc O. Ernst expands on the concept of WCS [5] between pairs by proposing a hypothetical soccer match

A. Batallones · K. Sanchez (✉) · B. Mott · D. Frank Hsu
Laboratory of Informatics and Data Mining, Department of
Computer and Information Science, Fordham University,
New York, NY, USA
e-mail: kisanchez@fordham.edu

A. Batallones
e-mail: abatallones@fordham.edu

B. Mott
e-mail: bmott290@gmail.com

D. Frank Hsu
e-mail: hsu@cis.fordham.edu

C. Coffran
Program for the Human Environment, The Rockefeller
University, New York, NY, USA
e-mail: cameron@rockefeller.edu

during which two referees determine whether the ball falls behind a goal line. Similar to Bahrami's proposal, Ernst's findings indicate that simply taking the approach of BF or a CF omits information which could lead to an optimal joint decision between the pair. However, while Ernst agrees that the WCS model can lead to a beneficial joint determination, his findings also indicate that there are improvements that can be made to the WCS model to achieve a more optimal joint decision. With Ernst's scenario, Bahrami's WCS model can be applied as the distance of the individual's decision (d_i) divided by the spread of the confidence distribution (σ), which is d_i/σ_i . A modified version of WCS (which closely resembles DSS) using sigma-square can produce a more accurate estimate through the joint opinion, which is represented as d_i/σ_i^2 . In an affirmation of Bahrami's research, Ernst also notes that joint decision-making comes with a cost when individuals with dissimilar judgments attempt to come to a consensus in such a manner. Bahrami and Ernst set forth very different experimental methods, but their aim is very much the same: to devise an algorithm for optimal decision-making between two people based on visual sensory input.

In the other direction, neural bases for decision-making and combining sensory information within senses have been studied by Gold and Shadlin [10] and Hillis et al. [1]. Koriati [11] indicated that there is no need to combine two heads' decisions under a normal environment. His suggestion is to simply take the decision of the most confident person.

Combinatorial Fusion Analysis (CFA), an emerging information fusion paradigm, was proposed for analyzing the combination of multiple scoring systems (MSS) (see Hsu et al. [12–14]). CFA has been shown to be useful in several research domains, including sensor feature selection and combination [15, 16], information retrieval, system selection and combination [12, 17], text categorization [18], protein structure prediction [19], image recognition [20], target tracking [21], ChIP-seq peak detection [22], and virtual screening [23]. These studies have shown in its respective domain that combination of MSS performs better than individual systems when the individual scoring systems perform relatively well and they are characteristically different [13, 14].

In a series of previous studies [24–26], a modified version of the soccer goal line decision proposed by Ernst is used as the data collection method. In this experiment, two subjects observe a small target being thrown into a grass field. The subjects are separately asked of their decision on their perceived landing point of the target and their respective confidences in their decisions. More recently, we conducted two sets of experiments with a total of 20 trials on two different days (12 trials and 8 trials) [27, 28]. In each of these trials, a small token was thrown into a grass field and landed at location $A = (A_x, A_y)$. Two

subjects P and Q standing 40 feet away from the landing site would perceive the landing site as at location $P = (P_x, P_y)$ and $Q = (Q_x, Q_y)$ with confidence radius σ_P and σ_Q , respectively. In these works, each visual cognition system is treated as a scoring system which assigns a score to each of the partitioned intervals in the common visual space. Then the problem of combining visual cognition systems is transformed to the problem of combining multiple scoring systems. The combination is analyzed using the CFA framework. Results obtained showed that combination by rank as well as by score can improve individual systems.

In this paper, we explore the issue of when and why a combination of two cognitive systems is better than each individual system using the CFA. In particular, we use the concept of "cognitive diversity" and the notion of "performance ratio" to analyze the outcome of the combination. Using the data set from the experiment with twelve trials [27], we demonstrate, as in other domain applications, that combination is positive (better than or equal to the best of the two individual systems) only if the two systems, based on weighted mean using confidence radius, are relatively good (higher performance ratio) and they are diverse (higher cognitive diversity).

Section 2 of this paper discusses two methods of combining visual cognition systems: statistical mean and combinatorial fusion. In Sect. 2.1, three statistical means M_1 , M_2 , and M_3 are calculated as average or weighted mean using the confidence radius as the weight. Based on these means, scoring systems p and q are constructed from the two visual cognition systems P and Q , respectively, in Sect. 2.2. Section 2.3 gives the method to combine these two visual scoring systems using the CFA framework. Section 3 gives the definition of cognitive diversity and the notion of performance ratio. Section 4 consists of examples, in particular the data set of an experiment with twelve trials of pairs of visual cognition systems [27]. Combination of these two visual cognition systems and analysis of the combination for the data set is discussed in more detail in Sect. 4.2 and 4.3. A summary of the results and possible future works is discussed in Sect. 5.

2 The CFA framework for combining two visual cognition systems

2.1 Computing various statistical means

When we make a decision based on visual input, we can consider this decision-making as a contemplation of various choices or candidates. Given two perceived locations $P = (P_x, P_y)$ and $Q = (Q_x, Q_y)$ (with confidence radius σ_P and σ_Q , respectively) of the actual landing site $A = (A_x, A_y)$, we wish to find a new location L (obtained by the joint decision of

P and Q) so that L is better than P and Q (distance between L and A is smaller than those between P and A , and Q and A). When determining a joint decision, typically an average or a weighted average approach is used to determine a mean. Average mean $M_1 = (M_{1x}, M_{1y})$ of the two locations $P = (P_x, P_y)$ and $Q = (Q_x, Q_y)$ is calculated as

$$M_1 = (P + Q) / 2, \tag{1}$$

and weighted means are obtained by

$$M_2 = (P/\sigma_P + Q/\sigma_Q) / (1/\sigma_P + 1/\sigma_Q), \tag{2}$$

and

$$M_3 = (P/\sigma_P^2 + Q/\sigma_Q^2) / (1/\sigma_P^2 + 1/\sigma_Q^2), \tag{3}$$

where P and Q are the perceived locations of the individual subjects P and Q , and σ_P and σ_Q are the confidence measurement of the two subjects, respectively.

2.2 Converting each visual cognition system to a scoring system

In the experiments we conducted, each of the two subjects provides an individually determined decision on where they respectively perceived the same target has landed in a field. Each coordinate on the field can be considered as a candidate for the respective participants' decisions of the perceived landing point. We are able to obtain a weight for each decision and their combination by asking each subject of a radius measurement of confidence around his or her decision. The smaller the radius measure of confidence, the more confident is the participant. We use radius R to calculate the spread (i.e., standard deviation) of the distribution around the perceived landing point, or σ . In our research, we use

$$\sigma = 0.5R. \tag{4}$$

2.2.1 Set common visual space

The σ values are used in Formulas (1), (2), and (3) to determine the positions of the means and denoted as M_1, M_2 , and M_3 respectively. The distance between M_i and A , $m_i = d(M_i, A)$, where A is the actual landing site, is used to evaluate the performance of M_i . With the field used as a two-dimensional coordinate grid, P , Q , and A are represented as x - and y -coordinates. Three formulas are used to calculate the mean of P and Q , as M_i , where $i = 1, 2$, or 3 . M_i falls somewhere in between points P and Q and is determined as a coordinate.

The longer of either segment PM_i or M_iQ is extended 30 % to the left to point P' or to the right to point Q' , respectively. The shorter side is extended more to create the widened observation area $P'Q'$ so that M_i is the midpoint of P' and Q' . We refer to the line segment $P'Q'$ as the common visual space (Fig. 1).



Fig. 1 The extension of PQ to $P'Q'$ based on M_i for $i = 1, 2$, or 3

We partition the length, $d(P',Q')$, of line segment $P'Q'$ into 127 intervals with midpoint d_i in each interval $i, i = 1, 2, \dots, 127$, and with each interval length $d(P',Q')/127$. The midpoint of the center interval, in this case, d_{64} , contains M_i .

2.2.2 Treat P and Q as two scoring systems p and q

Normal distribution probability curves for each participant are created with the point P and Q as the mean and using the confidence radii values, σ_P^2 and σ_Q^2 of P and Q as the variances of P and Q , respectively (see Fig. 2 in the case of 15 intervals). The following formula is used to determine normal distribution:

$$Y = (1/(\sigma\sqrt{2\pi})) * e^{[-(x-\mu)**2]/2\sigma**2}, \tag{5}$$

where x is a normal random variable, μ is the mean, and σ is the standard deviation. A normal distribution curve spans infinitely to the right and to the left. Therefore, our two scoring systems p and q create overlapping distributions that span the entire visual plane between P' and Q' . Scoring system p and scoring system q , respectively, scores each of the 127 intervals on the common visual space. For normal distribution functions with point P and Q as the mean and σ_P and σ_Q as the standard deviation respectively, each of the scoring systems p and q assigns interval d_i a score between 0 and 1 according to formula (5) (see Fig. 2 in the case of 15 intervals). These are the score functions s_p and s_q . The values of the score function s are sorted from highest to lowest to obtain the rank functions r_p and r_q , respectively (see Fig. 3). The d_i with the lowest integer as its rank has the highest score.

2.3 Combining scoring systems p and q using both score and rank combination

Let D be a set of candidates with $|D| = n$. Let $N = [1, n]$ be the set of integers from 1 to n and R be a set of real numbers. In the context of a CFA framework, a scoring

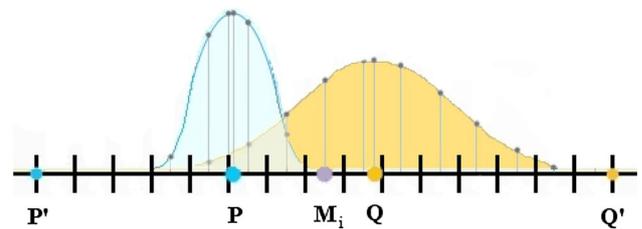


Fig. 2 Partition of $P'Q'$ into 15 intervals with center M_i

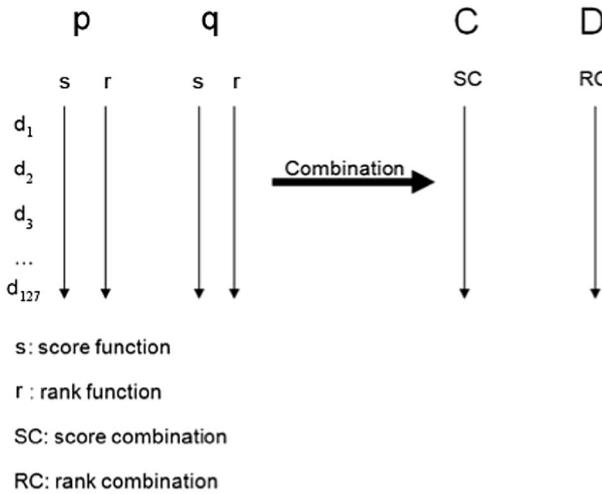


Fig. 3 Score and rank function for respective scoring systems p and q undergo CFA to produce score combination C and rank combination D

system A consists of a score function s_A and a rank function r_A on the set D of possible n positions (in this paper, $D = \{d_i | i = 1, 2, \dots, 127\}$).

In the setting of this paper, the score function s_C of the score combination of derived scoring systems p and q in our experiment is

$$s_C(d_i) = (s_p(d_i) + s_q(d_i)) / 2. \tag{6}$$

The score function s_D of the rank combination of the two scoring systems p and q in our experiment is

$$s_D(d_i) = (r_p(d_i) + r_q(d_i)) / 2. \tag{7}$$

When we sort $s_C(d_i)$ in descending order, we obtain the rank function of the score combination, called $r_C(d_i)$. When we sort $s_D(d_i)$ in ascending order, we obtain the rank function of the rank combination, called $r_D(d_i)$. The top ranked interval in $r_C(d_i)$ is called C . The top ranked interval in $r_D(d_i)$ is called D (see Fig. 3). These points are considered the optimal score and rank combination, respectively, and are used for evaluation of the combination result. The performance of the points (P , Q , M_i , C , and D) is determined by each respective point's distance from target A . A shorter distance indicates higher performance (Fig. 4).

3 Cognitive diversity and performance ratio

3.1 Cognitive diversity

Given the score function s_A of the system A and its derived rank function r_A , rank-score characteristic (RSC) function f_A , which is a composite function of s_A and the inverse of r_A , defined by Hsu et al. [13, 14] is a function from N to R and can be computed mathematically as (see Fig. 5).

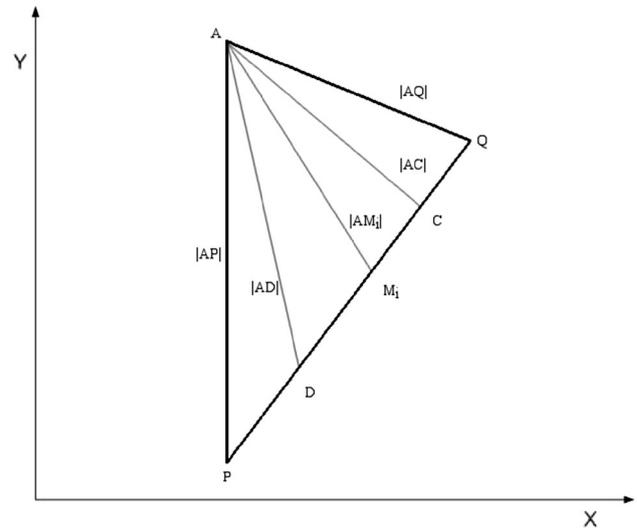


Fig. 4 Layout of M_i , $i = 1, 2$, or 3 , C , and D in relation to P , Q , and their distance to A . The distances between the 5 estimated points and A are noted on each line [24]

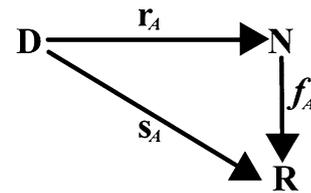


Fig. 5 Score function s_A , rank function r_A , and RSC function f_A of the scoring system A [13, 14]

$$f_A(i) = (s_A r_A^{-1})(i) = s_A(r_A^{-1}(i)). \tag{8}$$

The cognitive diversity between two scoring systems p and q , $d(p, q)$ is calculated using RSC functions f_p and f_q (also see [23]) as

$$d(p, q) = d(f_p, f_q) = \left(\sum_{i=1}^{127} (f_p(i) - f_q(i))^2 / 127 \right)^{1/2}. \tag{9}$$

3.2 Performance ratio

The performances of each P and Q for all trials are used in calculating the performance ratio. Performance of P (or Q) is determined by the distance between P (or Q) and A , $d(P, A)$ [or $d(Q, A)$], respectively. Shorter distance indicates high performances. Each distance is inverted and then multiplied by the maximum distance $md = \max\{d(P_i, A_i), d(Q_i, A_i) | i = 1, 2, \dots, 12\}$ for all trials. Let $MAX = \max\left\{ \frac{md}{d(P_i, A_i)}, \frac{md}{d(Q_i, A_i)} | i = 1, 2, \dots, 12 \right\}$. Then this set of numbers is each divided by MAX . In this way, the performance for each of the 12 P and Q is in the set $(0, 1]$. The smaller performance over the higher performance for

Table 1 Coordinates of P , Q , and A and confidence radius (σ) of P and Q for the 12 trials [27]

Trial	(P_x, P_y)	σ_P	(Q_x, Q_y)	σ_Q	(A_x, A_y)
1	(11.5, 134.5)	11.5	(78.5, 105)	16	(94, 124)
2	(23.5, 56)	7	(112, 96.75)	21.5	(28.5, 43)
3	(105, 134.25)	21	(78.5, 87.75)	22	(39.5, 119)
4	(229.25, 151.5)	14	(256, 162.5)	15.5	(216.25, 149.75)
5	(125.5, 13.5)	0.5	(112.75, 57.25)	3	(113.75, 46)
6	(184.5, 108.25)	21.5	(164.5, 249.75)	12	(173.25, 212.5)
7	(22, 190.5)	7	(17, 227.75)	6	(14.75, 195)
8	(98.75, 57)	12.5	(71.25, 25.5)	12	(16.5, 1)
9	(205.5, 15)	17	(204, 21.5)	6.5	(203, 26)
10	(100.5, 4.5)	19.5	(172, 25.25)	6	(127, 9.5)
11	(236.25, 43)	4	(234, 72.75)	4.5	(229, 51.5)
12	(98.5, -75.5)	10	(99, 30)	12	(96, 4)

P and Q is the performance ratio after it is normalized again among the twelve ratios to be in $(0, 1]$.

4 Example

4.1 Data set

We use the data set from an experiment of twelve trials conducted by the authors in [27]. Each trial consists of two volunteers P and Q with confidence radius σ_P and σ_Q . Each gives a visual cognitive estimate of the actual token landing site A as P and Q respectively.

Table 1 lists coordinates of $P (P_x, P_y)$, $Q (Q_x, Q_y)$, and $A (A_x, A_y)$ as well as the confidence radius σ_P and σ_Q of P and Q respectively.

4.2 Combination results and analysis

The decision of Participant p , marked as P , and the decision of Participant q , marked as Q , are used to obtain line segment PQ . The radii of confidence are used to calculate the two σ values to locate the coordinates of points M_1, M_2 , and M_3 along the extended $P'Q'$. To combine and compare the two visual decision systems of p and q , a common plane must be implemented to be evaluated by the different systems. The 127 intervals along the $P'Q'$ line serve as the common visual space to be scored.

When $P'Q'$ has been partitioned into the 127 intervals mapped according to M_i , the intervals are scored according to the normal distribution curves of P and Q using the standard deviation σ_P and σ_Q , respectively. Both systems assume the set of common interval midpoints $d_1, d_2, d_3, \dots, d_{127}$. Each scoring system, p and q , consists of a score function. We define score functions $s_P(d_i)$ and $s_Q(d_i)$ that map each interval, d_i , to a score in systems p and q ,

respectively. The rank function of each of the systems p and q maps each element d_i to a positive integer in N , where $N = \{x \mid 1 \leq x \leq 127\}$. We obtained the rank functions $r_P(d_i)$ and $r_Q(d_i)$ by sorting $s_P(d_i)$ and $s_Q(d_i)$ in descending order and assigning a rank value from 1 to 127 to each interval. C and D based on M_i , for $i = 1, 2$, and 3 , are calculated, and the distances to target A are computed. The point with the shorter distance from the target is considered the point with the better performance.

Table 2 lists the performance of (P, Q) , confidence radius of P, Q and performance of C and D based on M_i , $i = 1, 2$, and 3 . Table 3 lists performance for M_i , $i = 1, 2$, and 3 in the twelve trials. Table 4 gives comparisons of the performance of C or D to that of P and Q , and to M_i . We note that Koriat’s criterion, taking the decision of the most confident system, gives a correct prediction of 7 out of the 12 trials (Trials 1, 2, 4, 6, 8, 9, and 11). The score combination C or rank combination D obtained by CFA improves P and Q in 8, 7, and 6 out of the 12 trials when the common visual space mean is M_1, M_2 , and M_3 respectively. It is interesting to note that C or D improves P and Q in more trials based on M_1 than those based on M_2 or M_3 because M_1 does not take into consideration the confidence radius as weighted means (Table 4(a)). The same reason can be given to Table 4(b) where C or D can improve M_1 in more trials than M_2 or M_3 . In addition, in the 4 trials (Trials 3, 5, 10, and 12) that Koriat’s criterion fails to apply, they can all be improved using the CFA framework.

Figures 6 and 7 illustrate the performances of P, C, D, M_i , and Q for $i = 1, 2$, and 3 in Trials 2 and 7 respectively. In Trial 2, P performs quite good and has a higher confidence radius than Q . When given weighted means M_2 and M_3 , combinatorial fusion C or D performs better than P and Q . However, in Trial 7, P performs better but has a lower confidence radius than Q . In this case, C or D does not

Table 2 Performance of combination: (a) Performance of P, Q , (b) Confidence radius of P, Q , (c) Performance of C and D based on M_1, M_2 , and M_3 , respectively

Trial	(a) Per. (P, Q)	(b) Confidence Radius (σ_P, σ_Q)	(C)(1) Per. of C, D; based on M_1	(C)(2) Per. of C, D; based on M_2	(C)(3) Per. of C, D; based on M_3
1	(20.41 , 24.52)	(11.5 , 16)	(20.24 , 20.24)	(20.63, 20.07)	(20.14 , 20.14)
2	(13.93 , 99.3)	(7, 21.5)	(13.96, 13.96)	(13.91 , 13.91)	(13.91 , 13.91)
3	(67.25, 49.98)	(21 , 22)	(66.71, 49.94)	(66.72, 67.13)	(66.70, 67.15)
4	(13.12 , 41.74)	(14 , 15.5)	(14.47, 13.23)	(14.40, 13.11)	(14.48, 13.19)
5	(34.56, 11.29)	(0.5 , 3)	(34.38, 11.12)	(10.95 , 10.95)	(34.51, 10.94)
6	(104.86, 38.26)	(21.5, 12)	(37.70 , 37.70)	(37.63 , 37.63)	(37.95 , 37.95)
7	(8.53 , 32.83)	(7, 6)	(32.68, 8.44)	(32.88, 32.44)	(32.68, 32.68)
8	(99.5, 59.98)	(12.5, 12)	(60.13, 60.13)	(59.79 , 59.79)	(59.90 , 59.90)
9	(11.28, 4.61)	(17, 6.5)	(4.86, 4.64)	(4.86, 4.65)	(4.95, 4.56)
10	(26.97 , 47.68)	(19.5, 6)	(47.38, 26.68)	(47.73, 46.48)	(47.24, 47.24)
11	(11.17 , 21.83)	(4 , 4.5)	(11.08 , 11.08)	(11.08 , 11.08)	(11.22, 10.92)
12	(79.54, 26.17)	(10 , 12)	(79.12, 25.76)	(79.80, 78.53)	(78.86, 78.86)

Bold numbers indicate C and/or D perform better than P and Q in (C)(1), (C)(2) and (C)(3). Bold numbers indicate better performance of the two systems in (a) and higher confidence in (b)

Table 3 Performance of M_1, M_2, M_3 in 12 trials

	Trial											
	1	2	3	4	5	6	7	8	9	10	11	12
M_1	4.37	51.52	52.86	27.35	11.91	33.52	14.90	79.45	7.95	10.70	8.84	26.89
M_2	4.13	28.45	53.08	26.62	33.09	13.53	16.21	79.05	6.45	30.20	8.28	31.66
M_3	6.28	17.26	53.32	25.90	34.51	5.41	17.53	78.64	5.46	41.25	7.78	36.36

Each bold number indicates the performance of M_i in the Trial is better than P and Q . M_3 is best among M_i 's in Trials 2, 4, 6, 8, 9, and 11

Table 4 Comparisons of performance of C or D to that (a) of P and Q , (b) of M_i , and (c) of P, Q , and M_i (set of 36 cases in Table 2)

	(a) C or $D \geq P$ and Q	(b) C or $D \geq M_i$	(c) C or $D \geq P, Q, \& M_i$
M_1	1, 3, 5, 6, 7, 10, 11, 12 (8/12)	2, 3, 4, 5, 7, 8, 9, 12 (8/12)	3, 5, 7, 12 (4/12)
M_2	1, 2, 4, 5, 6, 8, 11 (7/12)	2, 4, 5, 8, 9 (5/12)	2, 4, 5, 8 (4/12)
M_3	1, 2, 5, 6, 8, 9 (6/12)	2, 4, 5, 8, 9 (5/12)	2, 5 (2/12)
Total	21/36	18/36	10/36

improve P and Q based on M_2 or M_3 when more weight is given to Q . Therefore, we observe that giving more weight to the better performer with a higher confidence leads to a combination which improves P and Q . We call such a case a positive case. In the following Sect. 4.3, we investigate in general when combination (either rank or score combination) can improve P and Q .

4.3 Positive cases versus Negative cases

We plot the result of a score or rank combination of P and Q , distinguishing positive cases as “□” or “◇” and

negative cases as “×” or “+” on the two-dimensional coordinate plane with the y-axis as the cognitive diversity $d(P, Q)$ and the x-axis as the performance ratio P_i/P_h (lower performance over higher performance) for all the trials for each $M_i, i = 1, 2, \text{ or } 3$. Each trial within each graph is noted as positive when rank or score combination performs better than both P and Q , and negative when it does not. The average for all positive cases and the average for all negative cases is also marked for each graph as “■” and “X” respectively.

Cognitive diversity between P and $Q, d(P, Q)$, is the diversity between two RSC functions f_p and $f_q, d(f_p, f_q)$, and

Fig. 6 Performance of *P*, *C*, *D*, and *Q* based on M_1 (a), M_2 (b), and M_3 (c) respectively for Trial 2, **a** Performance of *P*, *Q*, *C*, and *D* based on M_1 in Trial 2, **b** performance of *P*, *Q*, *C*, and *D* based on M_2 in Trial 2, **c** performance of *P*, *Q*, *C*, and *D* based on M_3 in Trial 2

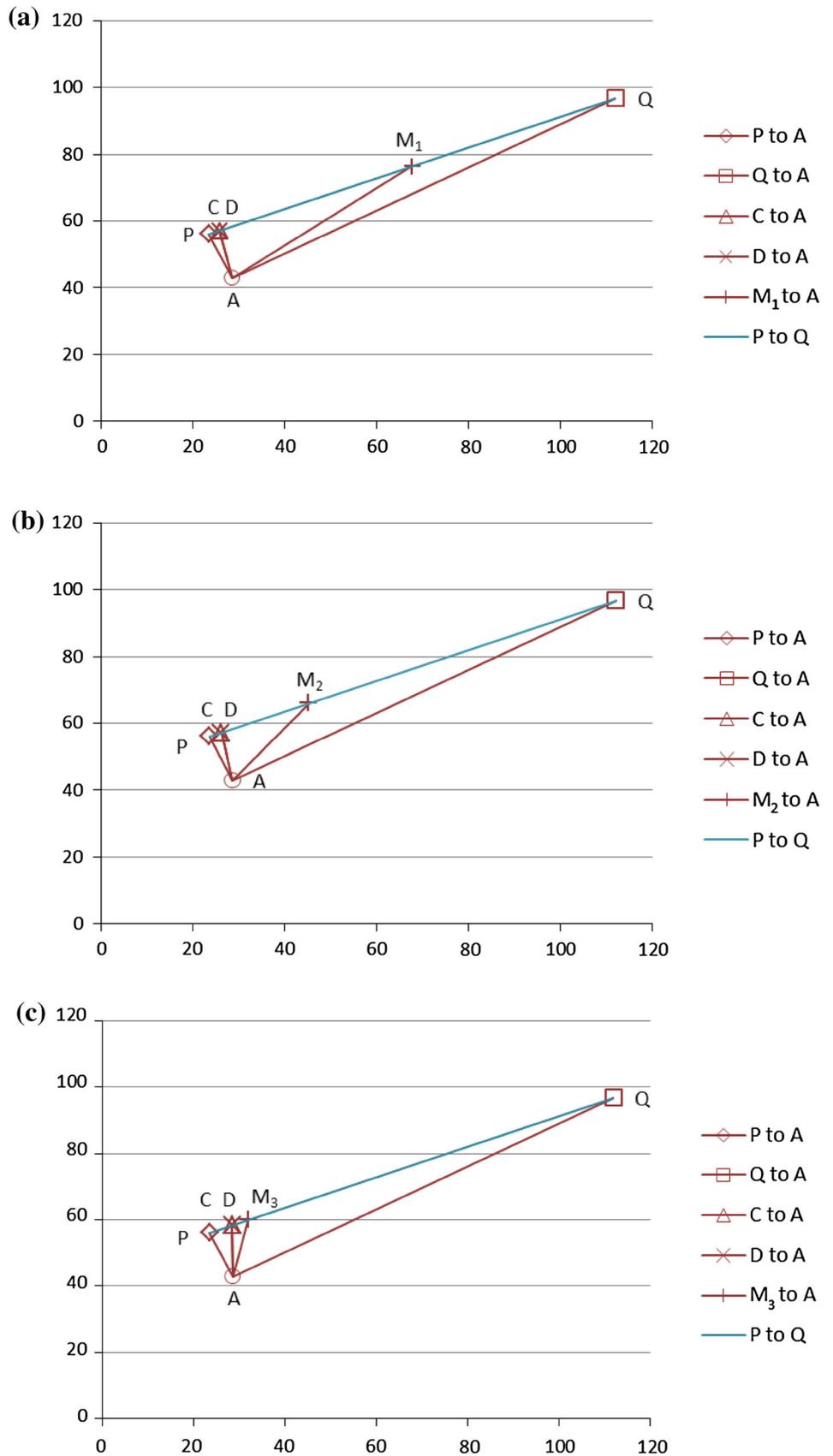


Fig. 7 Performance of *P*, *C*, *D*, and *Q* based on M_1 (a), M_2 (b), and M_3 (c) respectively for Trial 7, **a** Performance of *P*, *Q*, *C*, and *D* based on M_1 in Trial 7, **b** performance of *P*, *Q*, *C*, and *D* based on M_2 in Trial 7, **c** performance of *P*, *C*, *D*, and *Q* based on M_3 in Trial 7

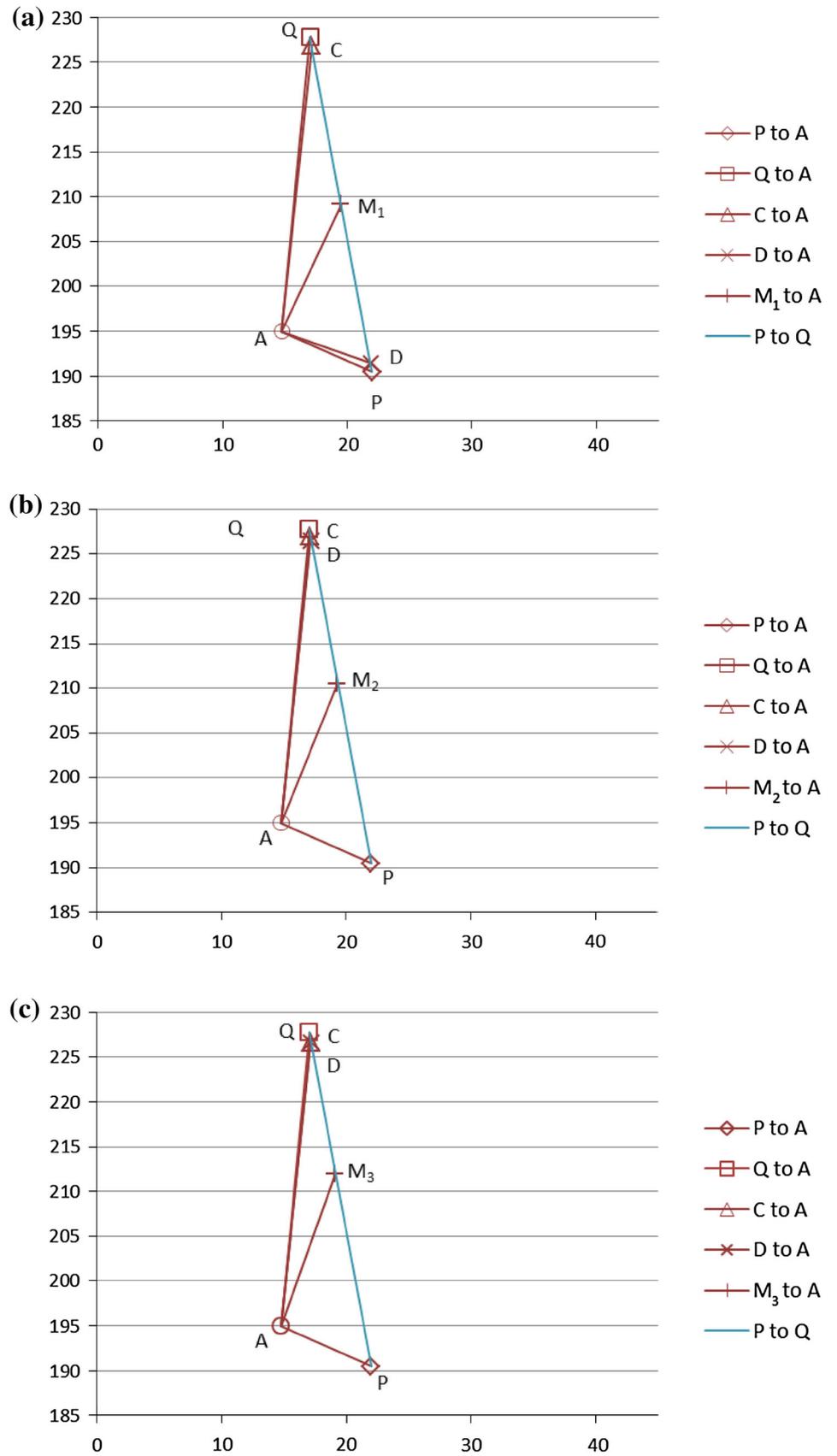


Table 5 Cognitive diversity

Trial	$d(p, q)$ in M_1	$d(p, q)$ in M_2	$d(p, q)$ in M_3
1	0.338434959	0.194412684	0.291268428
2	0.785773308	0.596314746	0.758254198
3	0.056297571	0.003988946	0.059975847
4	0.081718215	0.007480963	0.106193355
5	0.546649181	0.257373029	0.394914152
6	0.474573259	0.315880355	0.443491266
7	0.053300385	0.017943343	0.016196129
8	0.040005652	0.002402607	0.039874004
9	0.516003209	0.36502678	0.779226911
10	1	0.774343099	1
11	0.024840875	1	0.02068517
12	0.068319595	0.060741956	0.093857643

is calculated using formula (9). Cognitive diversity values are normalized to $(0, 1]$ in each case based on M_i , $i = 1, 2$, and 3 (see Table 5). Figure 8 depicts the positive versus negative cases based on each M_i , $i = 1, 2$, and 3 (Fig. 8a–c respectively) in terms of cognitive diversity (y-axis) and performance ratio (x-axis).

5 Summary and future work

In our previous work [27, 28], it has been demonstrated that combination of two visual cognition system using the CFA framework can improve each of the individual systems. In this paper, we analyze outcomes of these combinations according to positive cases or negative cases using the notions of cognitive diversity and performance ratio on the data set of an experiment with 12 trials [27]. It is demonstrated that in the majority of the 72 cases of rank combinations and score combinations ($12 \times 2 \times 3 = 72$) (see Fig. 8a–c), combination of two visual systems, based on weighted means M_2 or M_3 , can outperform each of the individual systems only if they each perform relatively well (with higher performance ratio) and they are diverse (with high cognitive diversity).

In an earlier work by Hsu and Taksa [12], it was shown that under certain conditions, rank combination can be better than score combination. In the current study, each of the six trials (Trials 1, 2, 5, 6, 9, and 10) has higher diversity than the remaining six trials. Similar to the results in [12], the six trials do have better rank combination (D) than score combination (C). It is also interesting to note that improvement in the other six

trials was carried out by rank combination only (Trial 3, 4, 7, 8, 11, and 12). In other cases, whenever score combination (C) improves P and Q , rank combination (D) can also improve. All these indicate that the CFA framework, which uses score and rank combination, is robust in analyzing combination and decision problems for visual cognition systems.

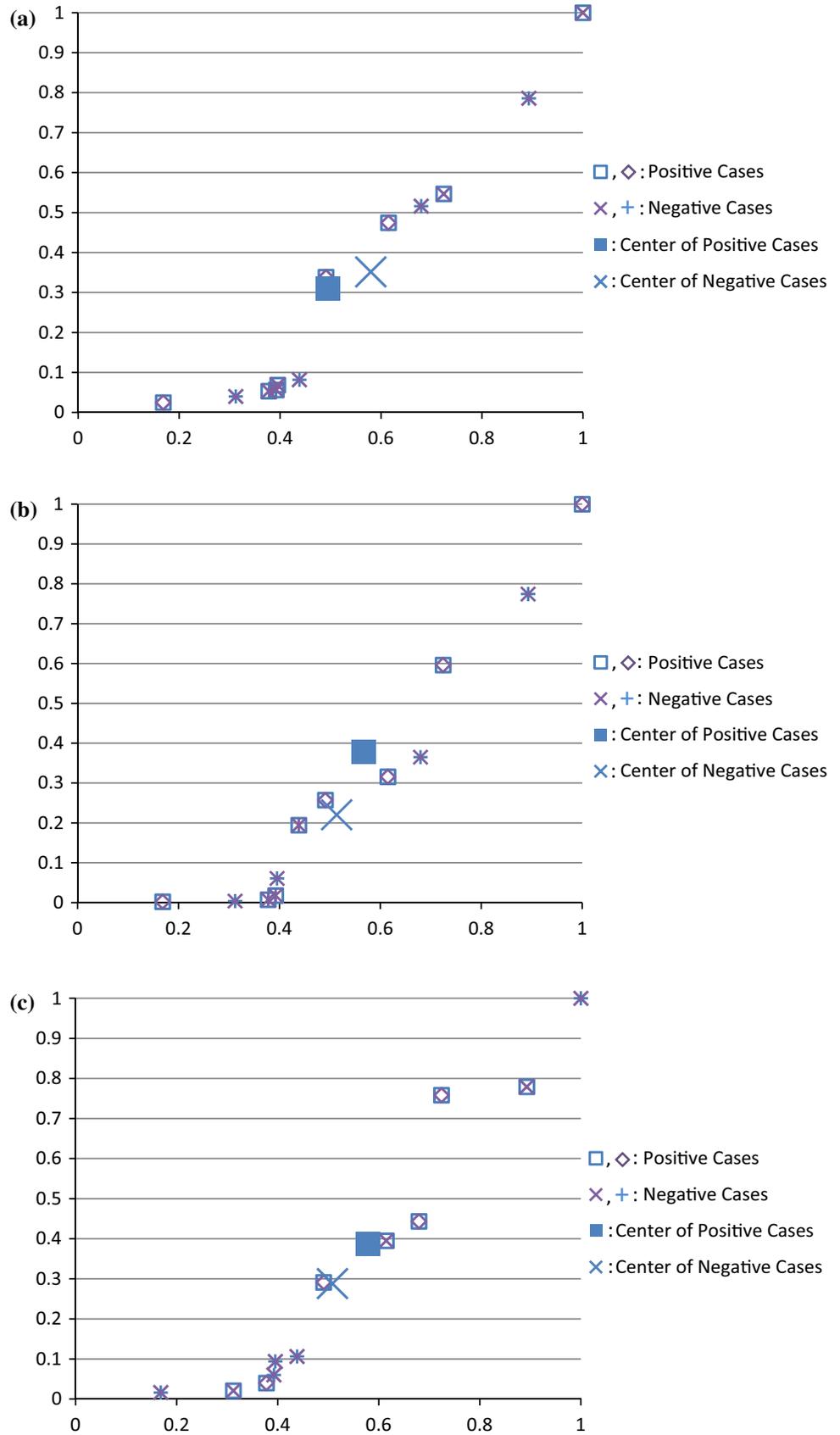
In the combination of decisions or visual cognition systems, as well as the integration of signals from different sensors, statistical means or weighted means such as M_1 , M_2 , or M_3 are often used [1, 3, 4, 5, 8]. It has been observed in these previous studies that M_3 , using $1/\sigma_P^2$ (or $1/\sigma_Q^2$) as the weight assigned to system P (or Q), provides better combination results. In our current study, when comparing M_1 , M_2 , and M_3 in each of the 12 trials, it is shown that M_3 is better than M_1 and M_2 in 6 of the 12 trials, while M_1 and M_2 are the best in 5 and 1 of the 12 trials respectively, independent of the performance of P and Q . So our current study supports that observation. However, when comparing improvements of M_i over P and Q , it was shown in our study that the statistical means M_1 , M_2 , and M_3 can improve P and Q in 4, 3, and 3 trials, respectively (see Table 3). On the other hand, the CFA framework (C or D) based on M_1 , M_2 , or M_3 can improve P and Q in 8, 7, or 6 trials. All these indicate that the CFA framework is a viable analytic method in combining visual cognition systems and can be generalized to analyze data in bioinformatics and neuroscience.

In summary, our CFA framework provides two criteria: performance ratio and cognitive diversity to guide us to combine two visual cognition systems with confidence radii. In the case of unsupervised learning or when the performance cannot be evaluated (e.g., the location of A is not known), cognitive diversity itself can be used to direct us when to combine (when the cognitive diversity is big enough) or how to combine (use rank combination or score combination) (see [12, 14, 21, 22, and 23]).

Our future work includes the following:

- (1) Apply CFA framework to the combination of more than two visual systems;
- (2) Study the effect of the number of partition intervals in the common visual space defined by $P'Q'$;
- (3) Use other diversity measurements such as Pearson's correlation (between two score functions s_A and s_B) and Kendall's tau (see [29]) or Spearman's rho (between two rank functions r_A and r_B); and
- (4) Apply CFA framework to combination of multiple sensing systems or combination of multi-modal physiological systems.

Fig. 8 Positive versus negative cases resulting from the 24 score and rank combinations in terms of cognitive diversity $d(P, Q)$ (y-axis) and performance ratio P_i/P_h (x-axis) based on M_1 (a), M_2 (b), and M_3 (c) respectively, **a** Positive versus negative cases based on M_1 , **b** positive versus negative cases based on M_2 , **c** positive versus negative cases based on M_3



Acknowledgments We thank two anonymous references for helpful comments and suggestions which led to improvement of the manuscript. DFH is partially supported by a travel fund provided by DI-MACS and CCICADA at Rutgers University.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Hillis JM, Ernst MO, Banks MS, Landy MS (2002) Combining sensory information: mandatory fusion within, but not between, senses. *Science* 298(5598):1627–1630
- Tong F, Meng M, Blake R (2006) Neural basis of binocular rivalry. *Trends Cogn Sci* 10(11):502–511
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433
- Ernst MO (2007) Learning to integrate arbitrary signals from vision and touch. *J Vis* 7(5):7
- Ernst MO (2010) Decisions made better. *Science* 329(5995):1022–1023
- Gepshtein S, Burge J, Ernst O, Banks S (2009) The combination of vision and touch depends on spatial proximity. *J Vis* 5(11):1013–1023
- Lunghi C, Binda P, Morrone C (2010) Touch disambiguates rivalrous perception at early stages of visual analysis. *Curr Biol* 20(4):R143–R144
- Bahrami B, Olsen K, Latham P, Roepstroff A, Rees G, Frith C (2010) Optimally interacting minds. *Science* 329(5995):1081–1085
- Kepecs A, Uchida N, Zariwala H, Mainen Z (2008) Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455:227–231
- Gold JJ, Shadlen N (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574
- Koriat A (2012) When are two heads better than one. *Science* 336:360–362
- Hsu DF, Taksa I (2005) Comparing rank and score combination methods for data fusion in information retrieval. *Inf Retrieval* 8(3):449–480
- Hsu DF, Chung YS, Kristal BS (2006) Combinatorial fusion analysis: methods and practice of combining multiple scoring systems. In: Hsu HH (ed) *Advanced data mining technologies in bioinformatics*. Idea Group Inc, Hershey, pp 1157–1181
- Hsu DF, Kristal BS, Schweikert C (2010) Rank-score characteristics (RSC) function and cognitive diversity. *Brain Inform* 6334:42–54
- Deng Y, Hsu DF, Wu Z, Chu CH (2012) Combining multiple sensor features for stress detection using combinatorial fusion. *J Interconnect Netw* 13(03n04)
- Deng Y, Wu Z, Chu CH, Zhang Q, Hsu DF (2013) Sensor feature selection and combination for stress identification using combinatorial fusion. *Int J Adv Rob Syst* 10
- Liu CY, Tang CY, Hsu DF (2013) Comparing system selection methods for the combinatorial fusion of multiple retrieval systems. *J Interconnect Netw* 14(01)
- Li Y, Hsu DF, Chung SM (2013) combination of multiple feature selection methods for text categorization by using combinatorial fusion analysis and rank-score characteristic. *Int J Artif Intell Tools* 22(02)
- Lin K-L, Lin C-Y, Huang C-D, Chang H-M, Yang C-Y, Lin C-T, Tang CY, Hsu DF (2007) Feature selection and combination criteria for improving accuracy in protein structure prediction. *IEEE Trans Nanobiosci* 6(2):186–196
- Liu H, Wu ZH, Zhang X, Hsu DF (2013) A skeleton pruning algorithm based on information fusion. *Pattern Recogn Lett* 34(10):1138–1145
- Lyons DM, Hsu DF (2009) Combining multiple scoring systems for target tracking using rank–score characteristics. *Inf Fusion* 10(2):124–136
- Schweikert C, Brown S, Tang Z, Smith PR, Hsu DF (2012) Combining multiple ChIP-seq peak detection systems using combinatorial fusion. *BMC Genom* 13(Suppl 8):S12
- Yang JM, Chen YF, Shen TW, Kristal BS, Hsu DF (2005) Consensus scoring criteria for improving enrichment in virtual screening. *J Chem Inf Model* 45:1134–1146
- McMunn-Coffran C, Paolercio E, Liu H, Tsai R, Hsu DF (2011) Joint decision making in visual cognition using Combinatorial Fusion Analysis. In: *Proceedings of the 10th IEEE international conference on cognitive informatics and cognitive computing*, 254–261
- McMunn-Coffran C, Paolercio E, Fei Y, Hsu DF (2012) Combining visual cognition systems for joint decision making using combinatorial fusion. In: *Proceedings of the 11th IEEE international conference on cognition informatics and cognition computing*, pp 313–322
- Paolercio E, McMunn-Coffran C, Mott B, Hsu DF, Schweikert C (2013) Fusion of two visual perception systems utilizing cognitive diversity. In: *Proceedings of the 12th IEEE international conference on cognitive informatics and cognitive computing*, pp 226–235
- Batallones A, McMunn-Coffran C, Mott B, Sanchez K, Hsu DF (2012) Comparative study of joint decision-making on two visual cognition systems using combinatorial fusion. *Active Media Technology. Lecture Notes in Computer Science, Series No. 7669*, pp 215–225
- Batallones A, McMunn-Coffran C, Mott B, Sanchez K, Hsu DF (2013) Combining two visual cognition systems using confidence radius and combinatorial fusion. *Brain and Health Informatics. Lecture Notes in Computer Science, Series No. 8211*, pp 72–81
- Ng KB, Kantor PB (2000) Predicting the effectiveness of naive data fusion on the basis of system characteristics. *J Am Soc Inform Sci* 51(12):1177–1189

Amy Batallones received her Bachelor of Science in Computer Science at Fordham University in New York. Her current research interests center around the methods of combination of visual cognition systems using informatics and information fusion.

Kilby Sanchez received a B.A. and an M.S., both in computer science, from Fordham University. His research interests include the application of combinatorial techniques in econometrics, economic forecasting, decision making, and financial modeling. He is also interested in the design of intelligent motion and signal detection systems using machine learning and data mining.

Brian Mott received a B.A. from the University of Buffalo and an M.S. from Fordham University. He is currently continuing his research in combinatorial fusion for financial decision makers. He is also interested in predictive analytics in the financial field.

Cameron Coffran received a B.S. and an M.S., both in computer science, from Fordham University. His current work involves clinical bioinformatics and software engineering for time series analysis. He

is also interested in further work related to combinatorial fusion and the management of large clinical data.

D. Frank Hsu Ph.D., is the Clavius Professor of Science and a Professor of Computer and Information Science at Fordham University in New York, NY. Dr. Hsu received an M.S. from the University of Texas at El Paso, Texas, and a Ph.D. from the University of Michigan. His research areas include combinatorics and algorithms; network interconnections and communications; and informatics and analytics. He has co-developed “combinatorial fusion algorithm” and

showed its applications in a variety of domain areas including cognitive science, bioinformatics, virtual screening, target tracking, and multimodal feature fusion. He has served on several editorial boards including Brain Informatics, Health Information Science (book monograph series), IEEE Transactions on Computer, International Journal of Foundation of Computer Science, Journal of Advanced Mathematics and Applications, and Journal of Interconnection Networks.