

Appendix for the COA (Coefficient of Agreement) Calculation and Errata Correction

February 26, 2013

1 Introduction

In our published paper [2], we extended the method [1, 3] to compute the COA (Coefficient of Agreement) based on the obtained pair-comparison user study data. The reasons why we cannot directly apply the methodology in the work of [1, 3] to compute the COA in our work [2], due to the following two reasons: (1) In our user study, the participants can have three options to choose (that is, besides “A over B” and “B over A”, an “undecided” option was added); by contrast, in the original papers [1, 3], only two options (“A over B” or “B over A”) are allowed. (2) In the original papers [1, 3], any two methods are compared via the pair-comparison methodology; by contrast, in our paper [2], we only used the pair-comparison methodology to compare our proposed method with other methods. In other words, we did not perform pair-comparisons between any two of “other methods”. Therefore, directly applying the COA calculation formula in [1, 3] will not work for our case. In our paper [2], we did an ad hoc tweaking on the COA calculation part, described in follow-up Section 2.

2 Tweaking of COA Calculation

In the work of [1, 3], the COA, u , is computed as follows:

$$p = \sum p_{ij} * (\sum p_{ij} - 1)/2, \quad (1)$$

$$u = 2 * p / ((s * s - 1)/2) * (t * (t - 1)/2) - 1, \quad (2)$$

where p_{ij} is the number of comparison pairs where our approach was preferred to the other method, s is the number of subjects, and t is the total number of methods in comparison (including ours).

In our paper [2], we tweaked the above equations to the following formulas:

$$t' = 2 * t, \tag{3}$$

$$u = 2 * p / ((s * s - 1) / 2) * (t' * (t' - 1) / 2). \tag{4}$$

Explanations for this ad hoc tweaking: (1) Since we provide 3 options for participants to select (not the 2 options in the original work [1, 3]), we multiply t by two (Eq. 3); (2) Since we only performed pair comparisons between our method and other methods (not like the work of [1, 3], any two methods are compared in a paired way), we did not subtract 1 as in the Eq. 2.

Note: This tweaking is purely ad hoc since it lacks rigorous mathematical proof (questions regarding this tweaking can be directly addressed to Xiaohan Ma, maxiaohan@gmail.com). Therefore, we do not recommend the above ad hoc tweaking for future similar COA analysis purpose. However, it is noteworthy to point out that this COA calculation is just a small and minor part of the whole paper. Its soundness is irrelevant to the main methodology, evaluation outcomes and conclusions of our work [2].

3 Errata Correction

We also found the two columns of Table 1 in the original paper [2] have errors (Table 2 and Table 3 are correct). For computing u and χ^2 in Table 1, we missed “multiplication by 2”. Therefore, the corrected u and χ^2 in Table 1 are as follows:

| # | u | χ^2 | p value | Motion Capture | Levine et al. '09 | Busso et al. '05 | Chuang et al. '05 |
|-----------|--------|----------|-----------|-----------------------|--------------------------|-------------------------|--------------------------|
| 1 | 0.074 | 67.400 | <0.001 | 9/10 | 8/10 | 11/7 | 13/7 |
| 2 | 0.069 | 65.199 | <0.001 | 5/2 | 11/7 | 12/7 | 11/8 |
| 3 | 0.072 | 66.400 | <0.001 | 5/11 | 9/7 | 14/6 | 11/8 |
| 4 | 0.073 | 67.199 | <0.001 | 8/10 | 9/7 | 12/7 | 12/6 |
| 5 | 0.082 | 71.800 | <0.001 | 10/9 | 8/10 | 14/6 | 11/9 |
| 6 | 0.082 | 72.000 | <0.001 | 7/10 | 8/10 | 15/4 | 12/7 |
| 7 | 0.076 | 68.800 | <0.001 | 8/9 | 11/9 | 12/6 | 11/8 |
| 8 | 0.086 | 74.000 | <0.001 | 8/9 | 10/7 | 12/7 | 14/5 |
| 9 | 0.100 | 81.599 | <0.001 | 10/8 | 12/7 | 12/3 | 14/5 |
| 10 | 0.0338 | 46.000 | <0.01 | 4/11 | 8/4 | 8/10 | 8/11 |

Table 1: Consistency and agreement test statistics for the head motion comparative evaluations. The number pair (e.g., X/Y) shown in each cell of the right part of the table denotes that the total number of the participants who voted for our approach is X and the total number of the participants who voted for the other comparative approach is Y.

References

- [1] M. G. Kendall and B. Babington-Smith. On the method of paired comparisons. *Biometrika*, 31:324–345, 1940.
- [2] B. H. Le, X. Ma, and Z. Deng. Live speech driven head-and-eye motion generators. *IEEE Transactions on Visualization and Computer Graphics*, 18(11):1902–1914, Nov. 2012.
- [3] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen. Evaluation of tone mapping operators using a high dynamic range display. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, pages 640–648, New York, NY, USA, 2005.