Colour Difference Modelling for Moving Images

Jin-Seo Kim[1], Maeng-Sub Cho[1], Bon-Ki Koo[1], M. R. Luo[2] and Stephen Westland[3],
[1]Digital Content Research Division, ETRI, Republic of Korea
[2]Colour & Imaging Group, University of Leeds, Leeds, UK
[3]Centre for Colour Design Technology, University of Leeds, Leeds, UK

Abstract

Psychophysical experiments were carried out to collect the observer accuracy data which are used to analyse the performance of state of the art colour difference formulae, and to derive a colour difference model for moving images. Three MPEG standard test streams were used to the study. The initial hypothesis for the data analysis of moving images is that the moving images are combinations of consecutive still images. Seven image quality attributes were asked to the observers. colour difference thresholds were analysed before the calculation of colour differences. The result showed that human visual system (HVS) is highly sensitive to the difference of memory colours especially to skin tone, and the sensitivity decrease along with the increase of spatial and temporal frequencies. Then the performance analysis of state of the art colour difference formulae such as CIELAB, CIE94, CIELAB CMC, CIEDE2000, SCIELAB and iCAM were performed to see which formula can best predict the differences of various image quality attributes for moving images. Wrong decision (WD) analysis was used and the results showed that CIELAB performed best for question of overall image difference. For sharpness difference, SCIELAB performed best followed by CIELAB. Colour difference models for moving images are proposed by use of temporal blur filter to the array of corresponding pixels throughout the sequences of images. CIELAB colour difference model with temporal blur performed best and it can be the final candidate of the novel colour difference model for moving images.

Introduction

Image quality may be affected by various attributes such as colour, sharpness, contrast, noise for still images. Furthermore, the image quality of moving images could be affected by temporal attributes such as movement and smoothness. State of the art colour difference formulae as well as colour appearance models provide the means of image quality prediction by calculating the pixel by pixel difference between original and test images. It is noted that all of the above formulae were developed for the target of still images and those have not been applied to the moving images.

The aim of this study is first, to verify which colour difference formula has the best performance, and second, to derive a model to predict the image quality of moving stream by analysing the difference distribution along the consecutive images in a single stream. To do that the hypothesis was first considered that was the moving images are combinations of consecutive still images. Pair-comparison experiment was carried out to collect the observer's perception of differences. Observer's performance, image difference threshold and performance of the seven colour difference formulae were calculated and analysed based on the observer's experimental data.

In this paper, psychophysical experiments for the image quality difference of human visual system are introduced follow by providing colour difference threshold analysis. Then the performances of seven state of the art colour difference models are presented. Finally, novel algorithm of colour difference calculation for moving images and its performance are introduced and conclude the paper with summary.

Experimental

To collect subjective response data for the various image quality attributes, it is required to perform psychophysical experiments which consist of light source, stimuli, and observer. Basically the same experimental data as the previous study was used in this study [1],[2]. However, only pair-comparison results were analysed in order to compare the performance of colour difference formulae. Figure 1 shows the test streams used in this study. The duration of the image play was confined to six seconds for all of the test streams in order to avoid observer fatigue during the psychophysical experiment.

***Figure 1** Test streams*

Lightness(L), Chroma(C), Contrast(CT), Noise(N), Sharpness(S), Compression(COM) and Temporal(T) are the image quality attributes used in this study. Six different levels of transform for each of seven attributes were applied to prepare test streams. The test streams were converted using SONY Vegas® 6.0 editing software. Total numbers of 129 which consist of 3 test streams × 6 rendering levels × 7 rendering attributes + 3 original streams were used to the experiments. The corresponding questions used in the psychophysical experiments are listed in Table 1.

***Table 1** Questions used in the experiments*

No.	Question
1	Do they look the same in overall quality?
2	Do they look the same in colour?
3	Do they look the same in sharpness?
4	Do they look the same in contrast?
5	Do they look the same in noise?
6	Do they look the same in movement?
7	Do they look the same in smoothness?

Colour Difference Thresholds

Colour-difference thresholds for each rendering attribute were calculated to see the perceptual thresholds of the human visual system against the different levels of rendering attributes. The first step to calculate the colour difference threshold is to calculate the z-scores for each rendering attribute and rendering level. Then calculate the colour difference between the original image and the corresponding test image by use of one of the standard colour difference formulae. Final step is to list up the z-scores and their colour difference values. Table 2 shows the list of z-scores and colour difference values in case of chroma rendering by use of CIELAB colour difference formula. In Table 2, from (a) to (g) corresponds to the results of questions 1 to 7 described in Table 1.

***Table 2** Colour Difference thresholds for each question-CIELAB*

	C	CT	COM	L	N	S	T
Mobil	3.5129	6.9139	10.1180	6.8435	11.6280	4.7993	0.0000
Susie	0.7052	1.2042	2.6606	2.8258	7.2500	0.9433	4.0000
Tennis	2.5444	4.3077	4.4109	5.1689	9.9138	3.2935	12.0000
mean	2.2542	4.1419	5.7298	4.9461	9.5973	3.0120	5.3333
STD	1.4262	2.8585	3.8997	2.0181	2.2061	1.9433	6.1101
CV	63.2683	69.0126	68.0601	40.8021	22.9867	64.5195	114.5644

(a) Overall

	C	CT	COM	L	N	S	T
Mobil	7.0976	6.4875	3.8516	6.9686	7.9258	2.2354	10.5180
Susie	1.5652	2.3019	2.3540	5.4123	6.5393	1.0477	6.2777
Tennis	3.6441	3.7508	2.7179	6.5398	7.7762	2.8576	10.6280
mean	4.1023	4.1801	2.9745	6.3069	7.4138	2.0469	9.1412
STD	2.7945	2.1256	0.7811	0.8039	0.7610	0.9196	2.4805
CV	68.1207	50.8500	26.2592	12.7458	10.2646	44.9243	27.1353

(b) Colour

	C	CT	COM	L	N	S	T
Mobil	2.0832	3.9854	4.5511	3.6719	10.4940	4.4654	11.2740
Susie	0.2868	1.2875	2.5493	0.9524	8.0330	1.0188	9.0976
Tennis	2.2842	2.8344	4.3183	1.9720	9.7734	2.5652	2.7713
mean	1.5514	2.7024	3.8062	2.1988	9.4335	2.6831	7.7143
STD	1.0998	1.3538	1.0947	1.3739	1.2652	1.7263	4.4169
CV	70.8893	50.0950	28.7618	62.4831	13.4121	64.3398	57.2562

(c) Sharpness

	C	CT	COM	L	N	S	T
Mobil	11.3430	10.5000	3.1972	19.5190	2.3631	6.4471	11.9310
Susie	4.1650	7.2936	2.8064	3.2337	2.0883	2.1224	0.6593
Tennis	7.8436	10.4540	2.4924	12.4190	3.1379	4.6696	6.8557
mean	7.7839	9.4159	2.8320	11.7239	2.5298	4.4130	6.4820
STD	3.5894	1.8381	0.3531	8.1649	0.5443	2.1737	5.6451
CV	46.1130	19.5211	12.4681	69.6430	21.5153	49.2572	87.0894

(d) Contrast

	C	CT	COM	L	N	S	T
Mobil	25.9420	34.5120	9.1649	17.8440	14.6800	5.2758	8.6127
Susie	31.2250	20.2440	4.6720	32.2750	61.5490	13.9460	3.6162
Tennis	6.9853	12.1300	3.9294	63.8720	16.9420	9.0118	6.7883
mean	21.3841	22.2953	5.9221	37.9970	31.0570	9.4112	6.3391
STD	12.7464	11.3311	2.8328	23.5415	26.4311	4.3489	2.5284
CV	59.6071	50.8229	47.8341	61.9561	85.1050	46.2096	39.8854

(e) Noise

	C	CT	COM	L	N	S	T
Mobil	23.1870	14.0820	4.1661	36.3340	11.2970	12.2720	8.9700
Susie	4.8432	7.1937	2.1238	12.7870	10.2510	3.2989	3.4918
Tennis	7.7512	10.4530	3.3470	11.2150	15.3450	4.1667	6.9636
mean	11.9271	10.5762	3.2123	20.1120	12.2977	6.5792	6.4751
STD	9.8591	3.4458	1.0278	14.0706	2.6904	4.9492	2.7716
CV	82.6614	32.5806	31.9955	69.9614	21.8773	75.2244	42.8033

(f) Movement

	C	CT	COM	L	N	S	T
Mobil	4.5085	4.8670	3.4721	5.7330	3.6199	2.9096	8.5295
Susie	1.0365	1.1864	3.9134	2.8367	2.4503	1.7824	3.7675
Tennis	2.4050	2.5134	2.6637	5.1423	3.7944	2.9693	6.9269
mean	2.6500	2.8556	3.3497	4.5707	3.2882	2.5538	6.4080
STD	1.7489	1.8640	0.6338	1.5304	0.7309	0.6687	2.4230
CV	65.9969	65.2756	18.9201	33.4837	22.2270	26.1844	37.8130

(g) Smoothness

For all of the questions except 'Noise' difference, 'Susie' test stream showed the lowest threshold for all of the six colour difference formulae. This means human visual system (HVS) has the highest sensitivity in discriminating the difference of skin tone. This results also support the fact that human perception of stimuli is highly related to the memory colour especially skin tone [3],[4],[5],[6],[7],[8]. The results also showed that 'Mobil' test stream had the highest threshold for all of the formulae and most of the questions except 'Noise' difference. The reason is 'Mobil' test stream does not include memory colour, moreover, it has high chromatic stimuli as well as high spatial frequency and relatively low temporal frequency. For the question of 'Noise' difference, 'Susie' test stream had the lowest threshold for 'temporal' rendering attribute whereas 'Tennis' test stream had the lowest threshold for 'Chroma', 'Contrast' and 'Compression' rendering attributes. And 'Mobil' test stream had the highest threshold for 'Lightness', 'Noise' and 'Sharpness-Blur' rendering attributes for the most colour difference formulae.

Overall CV values are slightly larger than that of still images [1] because observers are unfamiliar with this kind of experiments. The overall CV can be improved by the accumulation of observer experiences. One thing to be noted is that the CV values for 'Noise' rendering attribute had relatively lower than other rendering attributes in most questions except for the question of noise difference. This means that the sensitivity of the HVS in discriminating the noise difference is almost consistent with the change of observers.

From the data analysis, it can be found that HVS is highly sensitive to the difference of memory colours especially to skin tone, and the sensitivity decreases along with the increase of spatial and temporal frequencies. Also the sensitivity decreases when the stimuli are new to the observers. In this case, the sensitivity decreases even though the temporal frequency is relatively low.

Performance of Colour Difference Models – between original and rendered streams

Wrong decision (WD) analysis was used to test the performance of the current colour difference formulae. In this study, test stream was converted into the corresponding bitmap image sequences then the colour difference calculations for each bitmap images were performed. Finally the colour difference values for the bitmap sequences were averaged. Table 3 shows the results.

***Table 3** Performance results*

		CIELAB	CIE94	CMC	DE2000
Overall	ΔE_t	5.34	6.72	8.22	5.96
	WD	**27**	30	31	31
Colour	ΔE_t	7.12	6.72	8.56	5.96
	WD	38	40	40	39
Sharpness	ΔE_t	4.04	3.78	8.22	3.14
	WD	33	36	37	36
Contrast	ΔE_t	12.8	9.3	11.44	7.86
	WD	**25**	26	**25**	**25**
Noise	ΔE_t	10.84	9.06	12.26	8.06
	WD	10	10	10	10
Movement	ΔE_t	10.84	9.06	12.26	8.06
	WD	15	**14**	15	15
Smoothness	ΔE_t	1.96	1.78	2.1	2.22
	WD	23	22	22	**21**

		CAM02	SCIELAB	iCAM
Overall	ΔE_t	5.18	5.5	4.78
	WD	29	29	31
Colour	ΔE_t	7.36	2.64	2.0
	WD	40	38	**34**
Sharpness	ΔE_t	3.98	4.44	4.78
	WD	34	**31**	37
Contrast	ΔE_t	10.28	17.62	12.28
	WD	**25**	26	26
Noise	ΔE_t	11.12	10.92	7.44
	WD	10	10	10
Movement	ΔE_t	11.12	13.42	1.72
	WD	15	15	35
Smoothness	ΔE_t	2.16	1.98	0.84
	WD	23	24	24

In Table 3, CIELAB performed best followed by CIECAM02 and SCIELAB for question of overall image quality difference. For colour difference, iCAM performed best followed by CIELAB and SCIELAB. For sharpness difference, SCIELAB performed best followed by CIELAB. For contrast difference, CIELAB, CMC, CIEDE2000 and CIECAM02 performed equally best followed by CIE94, SCIELAB and iCAM. For noise difference, all of the formulae had the same scores. For movement difference, CIE94 performed best followed by other formulae except for iCAM. Finally for the smoothness difference, CIEDE2000 performed best followed by CIE94 and CMC. Overall performance of the formulae is slightly worse than the case of still image and the reason might be that the observers are not familiar to the experiments. However, similar to the case of still images, the colour difference formulae which consider spatial characteristics didn't perform better than those of conventional non-spatially considered formulae especially for the case of the spatially considered questions, e.g. noise, movement and smoothness.

Colour Difference Models for Moving Images

The state of the art colour difference models such as SCIELAB or iCAM colour difference calculations mainly consider spatial characteristics of the human visual system in discriminating image difference. Therefore they use contrast sensitivity function as spatial filter in order to throw away some spatial information by blurring the image. The purpose of the spatial filtering is to take into account the low sensitivity of human visual system for the discrimination of high spatial frequency in the image. In this study, the spatial filter is extended to temporal domain and a novel method of colour difference calculation model is proposed. The key of the algorithm is applying temporal blur through images under the assumption of the characteristic of human visual system that is some of the changes that take place between frames cannot be seen by the visual system because they are occurring too fast to be visible. So we must remove some of the temporal information by blurring the information in temporal domain.

The idea was implemented by first arranging each corresponding pixels throughout the test images. For example, if the test stream has six seconds duration, 525×380 image size, and 30 frames per second, then one pixel array has the size of 180(30×6), and the total number of array becomes 199,500(525×380). Once the pixel arrays are created, each pixel array is filtered by temporal CSF filter in the frequency domain and converted back to the spatial domain. Then the filtered pixel arrays are rearranged to the corresponding images. Figure 2 shows some temporally blurred images.

Figure 2 *Example of temporally blurred images*

It can be seen from the figure 2 that temporal blurring effect for each corresponding pixels may create image trail effects throughout the whole image sequences.

The total numbers of 180 modified images are now temporally blurred images. In this study, total numbers of 129 streams were temporally blurred and seven colour difference calculations were applied to complete the colour difference model for moving images.

Wrong decision (WD) analysis was again used for the performance test, and Table 4 shows the results.

Table 4 *Performance results*

		CIELAB	CIE94	CMC	DE2000
Overall	ΔE_t	3.18	2.52	5.86	2.44
	WD	**23**	25	26	24
Colour	ΔE_t	5.56	1.54	5.92	1.54
	WD	36	38	38	**37**
Sharpness	ΔE_t	3.18	2.52	3.26	2.44
	WD	25	26	29	27
Contrast	ΔE_t	8.72	8.7	9.26	7.82
	WD	**25**	26	**26**	**26**
Noise	ΔE_t	10.0	10.0	10.0	7.82
	WD	10	10	10	10
Movement	ΔE_t	10.0	8.7	10.34	10.34
	WD	15	**15**	15	15
Smoothness	ΔE_t	1.56	1.54	1.6	1.48
	WD	22	24	22	20

		CAM02	SCIELAB	iCAM
Overall	ΔE_t	3.22	2.64	2.2
	WD	24	24	25
Colour	ΔE_t	5.82	1.72	1.44
	WD	38	37	**32**
Sharpness	ΔE_t	3.22	2.64	2.2
	WD	25	**25**	31
Contrast	ΔE_t	8.68	10.08	6.68
	WD	**25**	26	26
Noise	ΔE_t	9.3	10.08	6.68
	WD	10	10	10
Movement	ΔE_t	9.3	10.08	2.08
	WD	15	15	38
Smoothness	ΔE_t	1.48	0.92	0.48
	WD	25	27	25

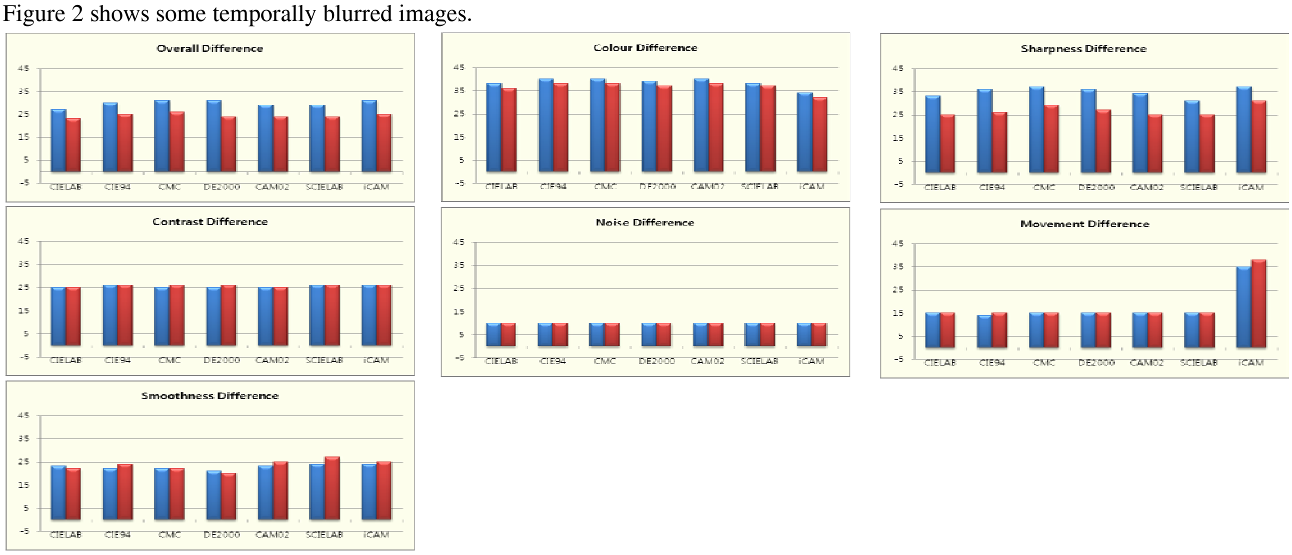

Figure 3 *Comparison of wrong decision results*

In Table 4, the ranking of the performance of seven colour difference models is the same as the previous case which is CIELAB performed best for overall image quality, iCAM performed best for colour, SCIELAB performed best for sharpness, CIELAB, CMC, CIEDE2000 and CIECAM02 performed equally best for contrast, all of the formulae had the same scores for noise, CIE94 performed best for movement, and CIEDE2000 performed best for smoothness.

Next, it is meaningful to compare the performance of state of the art colour difference models and proposed colour difference models, and Figure 3 shows the plots of wrong decision results for seven image quality attributes described in Table 1. In Figure 3, horizontal axis shows seven different colour difference models and vertical axis shows the WD values for each model. The blur bars correspond to the WD results of current colour difference models and red bars correspond to the WD results of colour difference models with proposed method.

It can be seen from figure 3 that in case of image difference prediction for overall, colour and sharpness, all of the image difference models using temporal blur modification performed better than the current colour difference models. For contrast, noise and movement, the performances of the models are almost the same except for the iCAM model in which current model performed better. And for the smoothness prediction, CIEDE2000 with temporal blur performed best but, the proposed colour difference model using CIE94, CIECAM02, SCIELAB and iCAM performed slightly worse than the current colour difference models.

In summary, image difference models using the proposed temporal blur algorithm enhanced the performance of difference prediction in case of overall, colour, and sharpness difference against current colour difference models. However, the proposed algorithm didn't enhance the performance in cases of contrast, noise, movement, and smoothness as expected. The possible reason is the accuracy of temporal contrast sensitivity function. Modifying the CSF may enhance the performance. Another reason of poor prediction is the inaccuracy of the experimental data. The observers participated in the experiments were the novices and they did not have the sufficient understandings of the concepts of movement or smoothness in the streams. Accumulated experiences of the experiments may also increase the accuracy of the experimental data.

It can be seen from the figure 3 that image difference model using CIELAB always performed better than just using CIELAB except for the cases of contrast and noise in which the performances are almost the same as others. This means that colour difference model using CIELAB is the best candidate for the colour difference model for moving images.

Conclusion

Performance tests for the six state of the art colour difference formulae were carried out by use of observer data. The analysis of colour difference threshold was carried out first, and 'Susie' test stream showed the lowest threshold for all of the questions except 'Noise' difference. This means that skin tone has the highest sensitivity to human visual system and the sensitivity decreases along with the increase of spatial and temporal frequencies. Then colour difference formulae were tested by the calculation of WD (%) method. For the question of overall image quality, CIELAB performed best followed by CIECAM02 and SCIELAB. And for sharpness difference, SCIELAB performed best followed by CIELAB. The complete analysis results will be included in the final paper.

A novel colour difference calculation algorithm for moving images was proposed. Temporal blur filter was applied to the algorithm. The performance of the proposed colour difference model was slightly better than state of the art colour difference models in cases of difference prediction of some image quality attributes. More study need to be performed to enhance the performance of proposed algorithm by modifying CSF. A new robust colour difference model for moving images will then be derived as a result of future study.

References

[1] J-S Kim, M-S Cho, M.R.Luo and S.Westland, Image Quality Modelling for Moving Streams, Proceedings of the fifteenth Colour Imaging Conference. IS&T/SID, Albuquerque, New Mexico, (2007)

[2] J-S Kim, M-S Cho, S.Westland and M.R.Luo, Image quality assessments for photographic images, AIC Colour 05, (2005)

[3] Stanley Coren, Lawrence M. Ward, and James T. Enns, *Sensation and perception*, Sixth edition, Wiley, pp. 114-115 (2004)

[4] Wichmann, F. A., Sharpe, L. T., & Gegenfurtner, K. R. Contributions of color to recognition memory for natural scenes. Journal of Experimental Psychology: Learning, Memory & Cognition, 28, pp. 509-520 (2002)

[5] Bodrogi P., and Tarczali T., Colour memory for various sky, skin, and plant colours: Effect of the image context, Color Research and Application, 26, pp. 278-289 (2001)

[6] Yendrikhovskij, S. N., Blommaert, F. J. J., & de-Ridder, H., Color representation and the naturalness constraint, Color Research and Application, 24, pp. 52-67 (1999)

[7] Yendrikhovskij, S. N., Blommaert, F. J. J., & de-Ridder, H., Representation of memory prototype for an object color, Color Research and Application, 24, pp. 393-410 (1999)

[8] Newhall, S. M., Burnham, R. W., & Clark, J. R., Comparison of successive with simultaneous color matching. Journal of the Optical Society of America, 47, pp. 43-56 (1957)

Author Biography

Jin-Seo, Kim received his M.S. degree in electrical engineering from Polytechnic University, Brooklyn, NY, USA, in 1993, and he is a Ph.D student in department of colour science, the University of Leeds, UK. He is a senior research engineer at ETRI since 1993. His research interest includes colour science, colour reproduction, colour management system, and digital cinema.