EG UK Theory and Practice of Computer Graphics (2005)
L. Lever, M. McDerby (Editors)

Towards Realism in Facial Image Prototyping: Results of a Wavelet MRF Method

Bernard Tiddeman[1], Michael Stirrat[2] and David Perrett[2]
1. School of Computer Science, University of St Andrews
2. School of Psychology, University of St Andrews

Abstract

The ability to combine multiple images to produce a composite that is representative of the set has applications in psychology research, medical imaging and entertainment. Current techniques using a combination of image warping and blending suffer from a lack of realism due to unrealistic or inappropriate textures in the output images. This paper describes a new method for improving the representation of textures when blending multiple facial images. We select the *most likely* value for each pixel, given the values of the neighbouring pixels, by learning from the corresponding values in the training set i.e. we use a Markov Random Field (MRF) texture model. We use a multi-scale neighbourhood and separate low and high frequency information using a wavelet transform. This ensures proper correlations of values across spatial scales and allows us to bias the global appearance to the mean for the set, while selecting more specific texture components at higher resolutions. We validate our results using perceptual testing that shows that the new prototypes improve realism over previous techniques.

Categories and Subject Descriptors (according to ACM CCS): I.4.7 [Image Processing and Computer Vision] Feature Measurement (Texture), I.3.8 [Computer Graphics] Applications.

1 Introduction

The ability to capture the essential characteristics of a set of images has found application in facial perception research, medicine and entertainment. In facial perception research, early theories of facial attraction focused on the idea that averageness was attractive. This was based on evidence from other species, where increased survival rates are correlated with being closer to the mean (e.g. in body size and shape). Photographic, and later digital, blending of human face images supported this idea, with the blends being rated as considerably more attractive than the original faces. Perrett et al [PMY94] showed that attractive faces differ systematically from average, by constructing averages of a more attractive subset. The researchers then began to investigate the role of other attributes, such as age [PPL*02], health [JLB*04], masculinity/femininity [PPC*] and familial resemblance [DEB04] in facial perception. In particular they are interested in the mechanisms by which we learn to find certain faces attractive, and the underlying evolutionary and neurological processes involved.

In addition to producing prototype faces, methods were developed to transform images between sets, e.g. to age or masculinise an individual's face image. The methods are based on adding the differences between two prototypes to an individual's face image [RP95]. This has applications in other areas, such as entertainment, e.g. for digital ageing of actors, and medicine, e.g. for predicting the outcome of surgery by learning from examples. Prototyping itself has found application in other areas. For example for constructing average medical images for facial surgery planning (e.g. [TDR00]) and brain analysis (e.g. [TT96]).

The usual method for creating prototype images is first to spatially align the component images to the mean shape of the set, and then to find the average colour of each pixel in the output image. This method does not properly align the fine detail textures and so the output images have blurred (i.e. very smooth), and therefore inappropriate textures. Hence improvements to image prototyping have focussed on improving the representation of textures.

In the remainder of this paper we first review the relevant literature from the image prototyping, image fusion and texture synthesis literature. We then describe our new technique for synthesising prototype facial images and present the results of a perceptual experiment to demonstrate the effectiveness of the new algorithm.

2 Literature Review

2.1 Facial Image Prototyping Literature

The creation of prototype images has a long history dating back to the methods of Francis Galton who created photographic averages by using multiple exposures after aligning the eyes and mouth [GAL78]. More recently, prototypes have been created by digitally blending faces together, after normalising the shape to the average using image warping [BP93]. Simple averaging of the spatially aligned images does not produce representative textures and several improvements have been suggested. Wavelet-magnitude based texture prototyping [TBP01] uses the smoothed magnitude of the wavelets as a measure of the average local activity in different spatial locations, orientations and scales across the set. The wavelet values of the simple average are locally rescaled to approximate these activity levels. An alternative wavelet histogram method has been proposed in which the histograms of the prototype image and its wavelet subbands are modified to match the (mean) histograms of the image set [MCV04]. These two methods produce visually similar results, with random textures being well represented, but more

structured textures (such as the hair) still appearing rather unrealistic and unrepresentative of the set.

An alternative approach is to use a local neighbourhood surrounding each point to estimate the most likely pixel value for the prototype [TID04]. By comparing the values in the neighbourhood with the matching neighbourhoods in the training set a probability distribution can be estimated. The highest probability value is selected from this distribution, then this pixel forms part of the neighbourhood for succeeding pixels. A causal neighbourhood, spanning two neighbouring scales, was used for efficiency and reconstruction stability. The probability distributions were estimated by smoothing the histograms of neighbourhood values. Certain choices of smoothing parameters caused locking of the algorithm onto large parts of a single image in the training set. To avoid this behaviour the histogram smoothing was increased at lower spatial scales, to give a more average global appearance, but leaving more specific (and therefore realistic) fine scale textures. Although no perceptual study was conducted, the results appear very realistic, although there is a problem of occasional discontinuities in the reconstructed images. These discontinuities are probably caused by the overlap in information content between neighbouring scales in a multiresolution (Gaussian) pyramid.

2.2 Image Fusion Literature

Image prototyping is a specialised example of multiple image fusion. The more usual problem is that of fusing multiple images of the same scene, taken using several different sensors (e.g. visible and infra-red) or taken under different conditions (e.g. focused on different objects or with different exposures). Methods based on wavelet pyramids have proved successful, with various algorithms for combining the individual wavelet components proposed. Point-based methods include selecting the wavelet coefficient with the largest absolute value or making a simple average. Other methods inspect the values in a window about each point when calculating the value. These methods include choosing the pixel from the image with the largest absolute value in the window [LMM95]), making an average of the matching points that is weighted by the local activity levels in the window [BK93] or weighting the samples based on the contrast sensitivity of the human visual system [WRM95]. Several comparisons of wavelet-based image fusion schemes have been conducted [ZB99] [BB04] [HCB02] and these indicate that over-complete wavelet decompositions are preferable to critically sampled decompositions (for reconstruction stability and fewer artefacts), that region-based methods give improved results over point-based methods and shorter filters reduce the number of artefacts in the resulting fused images.

2.3 Texture synthesis by analysis

The main problem in previous facial prototyping methods has been identified as the lack of appropriate facial textures in the output images. Synthesising patches of texture from examples has been the focus of a considerable number of research papers. The statistical nature of textures inspired methods that used global properties of the images, such as the image histogram and the histograms of wavelet subbands [HB95]. Starting with random noise, the image and subband histograms are alternately matched to the target histograms in an iterative process. This method produced excellent synthetic results for random textures, but was not able to reproduce more structured textures, such as hair. Extensions to the method using additional measures of the wavelet histograms (such as correlations within and across subbands) improved the results, but still failed on more complex textures [SP98].

An alternative to global optimisation of texture parameters is to look at local features. In MRF-based texture synthesis [JUL62] [HS81] [CC85a] [CC85b] [CJ83] it is assumed that the probability distribution of a pixel's intensity is dependent on its neighbours. The aim is to construct a texture such that the local conditional probability distribution functions of the synthesised image match those of the original texture sample. The original MRF methods proved to be very slow due to the need to rebuild the histograms or select suitable matching neighbourhoods at each pixel by scanning the entire image. Several methods that approximate the full MRF method have been devised, usually by searching the example image for the best matching n pixels, and choosing one of these using stochastic sampling [EL99]. Speed optimisations include using a multiscale approach [PL98] (i.e. building up a low resolution approximation to the texture and then refining it to higher resolutions) and fast search algorithms [WL00]. Performing the texture synthesis in the wavelet domain has allowed proper separation of information at different spatial scales [ZWT98], including a very efficient approximation using a neighbourhood consisting of only lower resolution subbands [DEB97].

3 Method

The prototyping method starts in the same way as the previous techniques [BP93][TBP01][MCV04], by using feature-based image warping [WOL90][RM95] to normalise the shape of the example faces to the average shape for the set (Figure 1). The facial feature points can be labelled by hand or placed automatically e.g. using active shape models [CTL95]. Alternatively optical flow can be used to align the component images automatically [BP96][VP97].

Each example image is then transformed into a wavelet basis. There is a wide possible choice, not only of particular wavelet filters (e.g. orthogonal or bi-orthogonal) to use, but also of how they are applied. The wavelet subbands can be critically sampled so the transformed image pyramid contains the same amount of data as the original, or can be sampled at the original resolution leading to an over-complete (i.e. redundant) representation. There is also a choice between subsampling the low-pass filtered image or leaving it at the original scale. In this work we use the results from the image-fusion literature to chose a redundant wavelet basis (sub-sampling only the low-pass filtered image at each level of the wavelet pyramid) and we use short filters to

reduce reconstruction artefacts. The decomposition produces horizontally and vertically filtered components, and so includes some directional information and we use an exact reconstruction algorithm.

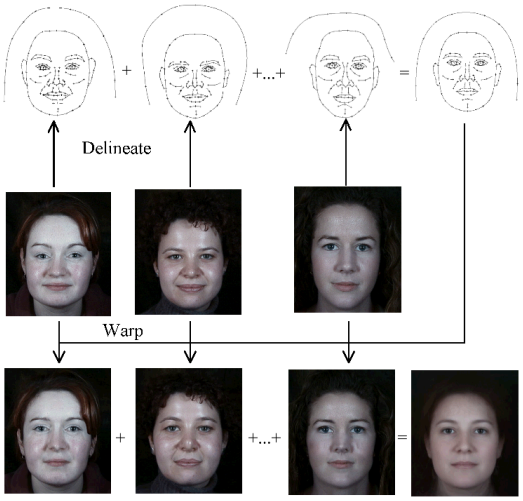

Figure 1: *The wavelet-MRF textured blending algorithm is an extension to the untextured algorithm illustrated above. The input faces (Centre row) are delineated (top row) and the average shape is found as the mean set of delineated feature points. The input images are all warped into this shape (bottom row) and the colours are averaged at each pixel to produce the average (bottom right).*

The filters we use are based on a first derivative of Gaussian approximation, with the low-pass residual subsampled by a factor of 2 in both x and y during the construction of each level of the wavelet pyramid. This is for greater efficiency, particularly when processing the lower resolutions. As with the filters used in [TBP01], we can perform an exact reconstruction by up-sampling, filtering and addition. The down and up-sampling of the low-pass filtered image requires that the high-pass components are convolved with different filters at odd and even pixels. Figure 2 shows the analysis and synthesis process and Table 1 gives the corresponding filter coefficients.

The re-texturing algorithm starts by making a low-resolution approximation to the prototype image by blending the low-pass residuals from the input images. Each successive finer resolution subband is then built up by scanning across the image and picking the most likely wavelet coefficients at each point from the conditional probability distribution. The conditional probability density is estimated by sampling from a fixed location across the spatially aligned example images. The probability distribution is estimated from the example images by smoothing the histogram of example values with a Gaussian function i.e. we use a Parzen window method.

	-2	*-1*	*0*	*1*	*2*
H		0.25	0.5	0.25	
G			1	-1	
F		0.5	1	0.5	
L_1		-0.25	0.25		
L_2	-0.125	-0.375	0.375	0.125	
K_1		0.125	0.75	0.125	
K_2	0.0625	0.125	0.625	0.125	0.0625

Table 1: *The wavelet filters used.*

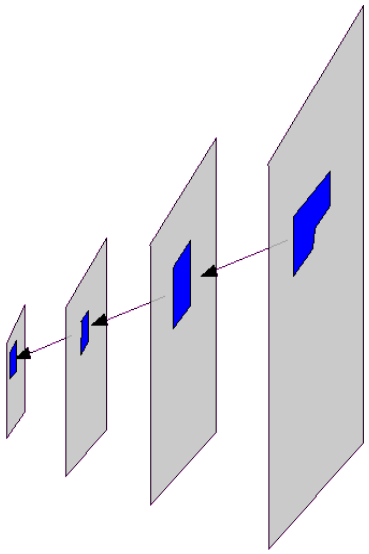

Figure 3: *The distribution at each point is conditional on the values of the pixels in its multi-resolution neighbourhood. The pixel's neighbourhood consists of a 12 point non-symmetrical half plane at the current resolution, a 3 by 3 block at the next lower resolution and a single pixel from each of the remaining lower resolution subbands.*

The neighbourhood of pixels we use includes a 12 pixel non-symmetrical half plane (NSHP) neighbourhood at the current resolution, a 3 by 3 symmetrical neighbourhood at the preceding resolution, and 1 pixel from each of the preceding resolutions (Figure 3). The use of symmetrical low-res neighbourhoods in addition to the NSHP high-res neighbourhood helps to stabilise the reconstruction, without the need for optimising the probabilities of all the pixels simultaneously. We do not assume total independence between subbands, but use all the information available from previous points in the neighbourhood to estimate each of the coefficients at the current point.

An algorithm for estimating the conditional distribution is given below:

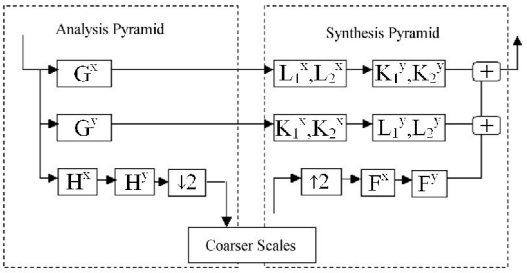

Figure 2: *The construction of one level of the wavelet pyramid. A pair of filters in a box indicates application to even and odd pixels respectively.*

Algorithm 1: Calculate conditional distribution
inputs:
input image's wavelet transform (WT) ***I****,*
array of example images' WTs ***J****,*
location ***(x,y)****,*
subband ***s***
neighbourhood ***N****,*
smoothing parameter ***h****,*
example sample size ***M****,*
histogram bin width ***bw***
number of bins ***bcount***
variables:
Float histogram array ***p*** *of length* ***bcount***
Float array ***u*** *for neighbourhood in* ***I***
Float array ***v*** *for neighbourhood in* ***S***[***i***]
begin:
1. *Initialise* ***p*** *to* ***0***
2. ***u*** *= values in* ***N*** *of* $\boldsymbol{I}_s(\boldsymbol{x},\boldsymbol{y})$
3. *for* ***i*** *= 0 to* ***M***
   3.1 ***v*** *= values in* ***N*** *of* $\boldsymbol{J}_s[\boldsymbol{i}](\boldsymbol{x},\boldsymbol{y})$
   3.2 $\boldsymbol{p}[\boldsymbol{J}_s[\boldsymbol{i}](\boldsymbol{x},\boldsymbol{y})/\boldsymbol{bw}]$ += *Gaussian*(***v***, ***u***, ***h***)
4. *Smooth* ***p*** *with 1D Gaussian of width* ***h*** *and re-normalise*
5. *Return probability distribution* ***p***

The function *Gaussian*(***v***, ***u***, ***h***) returns the value of a uniform multidimensional Gaussian of standard deviation ***h*** centred on ***u*** and evaluated at ***v***.

The non-parametric distribution is estimated by smoothing the N-dimensional histogram using a multi-dimensional Gaussian as the kernel function. The shape of the distribution is critically dependent on the kernel's smoothing parameter, in this case the width of the Gaussian. The "optimal" value, h_0, given in [PL98] for the smoothing parameter is only optimal if the true underlying distribution is Gaussian [SIL86]. Experiments using the Gaussian assumption for single texture synthesis in the spatial domain proved unsuccessful because the distribution is smoothed too much [TID04]. This may have been due to the high correlations between scales leading to an elongated shape for the Gaussian, and does not appear to be a problem in the wavelet domain [ZWT98].

Another problem identified in the spatial MRF work was that the algorithm often "locks" on to a single image at a low resolution, resulting in a final image that is just a patchwork of areas copied from a small number of images in the sample [TID04]. The same problem can occur in the wavelet domain. Increasing the value of the smoothing parameter, *h*, prevents locking but leads to a loss of apparent texture, as shown in Figure 4. To avoid this problem the width of the smoothing window can be increased linearly with increasing scale. We have experimented with different values of *h* and have found $h=(1+0.5l)h_0$, for pyramid level *l*, produces good results for a range of image sets of size 640 by 480 pixels. This keeps the distribution similar to the Gaussian mean at low resolutions, but retains the high-resolution textures. Further experimentation will be required to find the perceptually optimal value for *h* and the relationship with image resolution.

Figure 4: *The effect of using different values of the distribution smoothing parameter, h. Setting the value of h too high ($h=3h_0$) textures are lost (top left) giving an output similar to the untextured prototype (bottom left). Setting the value too low ($h=2h_0$) (centre top) the texture "locks" on to an individual face's textures (centre bottom). The function $h = (1.0+l*0.5)h_0$ for level l in the wavelet pyramid produced both realistic high-res textures and average global appearance (top right), but setting the gradient too high (e.g. $h=(1+l)h_0$) and the image appears noisy, rather than well textured (bottom right).*

4 Results

4.1 Visual evaluation

Figure 5 shows the results of blending several different sets of face images, grouped by approximate age, sex and ethnic background, and compares two previously published methods; wavelet-magnitude (WM) re-texturing [TBP01] and spatial-domain MRF (spatial) [TID04,] with the new approach. The WM method simply magnifies the wavelet coefficients to have the average magnitude value for the set at each point in each subband. The spatial method performs an MRF algorithm on a multiscale pyramid. Visual inspection of these prototypes indicates that the new wavelet MRF method provides a small

increase in realism in addition to retaining group characteristics such as age and gender. Although it is sometimes possible to identify some textural features from individuals in the blends, the transitions between texture components is not obvious and the end result looks like a distinct individual of "typical" appearance. Figure 6 gives further examples of the new prototypes.

Figure 5: *Comparison of the 3 different prototyping methods. The images with the highest (left) and lowest (right) mean rating for realism from each set are shown. Top, the new method (MRF): ratings left 4.77 right 2.69. Centre, wavelet magnitude method (WM): ratings left 4.48 right 2.69. Bottom, spatial MRF (spatial): ratings left 2.77 right 1.48. All on a 7 point (0-6) scale.*

4.2 Experimental evaluation

In order to test the new method we conducted a web-based experiment. We used two experiments, in the first subjects were asked to estimate the *age* of the faces, and in the second subjects were asked to rate the *realism* of the images produced. Java applets were used to present the images to subjects in a (pseudo) randomised order. For realism subjects where asked to rate the faces on a 7 point scale from 0 (very unrealistic) to 6 (very realistic). We compared the new method (MRF) with the WM and spatial methods. Each subject was shown male and female prototypes in the approximate age groups 8-12, 13-18, 25-45 and over 45 years old created by each of the 3 methods (spatial, WM and MRF), plus another set created by a variant on the method described here (results not shown).

Figure 6: *Further examples of wavelet-MRF prototypes of different groups of faces.*

The web-based age-rating experiment had over 500 responses. As with all web-based experiments, some filtering of the data was required to remove participants who did not complete the experiments in a sensible fashion e.g. by just clicking through after becoming bored. Participant scores were removed from analysis when the IP address was duplicated or where unreasonably young ages were entered (e.g. the age rating was less than 10 years for wavelet magnitude prototypes for age bracket greater than 13 years). This left 488 participants. In order to compare the age ratings we used the WM method as the "ground truth" and compared the other methods with it. This method has previously been shown to produce prototypes that are not perceived as significantly different from the mean age of the set [TBP01].

We calculated the mean and standard deviation for each rated image across all participants. There was a high degree of concordance of age estimates across participants (Chronbach's Alpha = 0.758). We then used a paired samples t-test to compare the age differences, but neither method was significantly different from the WM method at the 0.05 confidence level (MRF: t_7=1.53, p=0.164, Spatial: t_7=1.8, p=0.102). The spatial method tended to produce images a little older than the WM method, and the MRF method a little younger (Figure 7).

For the realism experiment we removed participant scores from the analysis when the IP address was duplicated, if they had been rejected in the age rating experiment or when the standard deviation of the ratings was less than 1.0 and when they used <=3 categories of the realism scale. This left 367 participants. From these we calculated the mean and standard deviations of the perceived realism ratings for each prototype. There was a high degree of

concordance of estimates across participants (Chronbach's Alpha = 0.849). The results of the perceptual study show that the new prototypes do indeed show a significant increase in realism (Figure 8). The mean ratings were spatial: 1.6, WM: 3.3 and MRF 4.0. We used a paired t-test to compare the mean ratings for the 8 prototypes. The MRF method was significantly more realistic than both the spatial method (t_7=8.08, p<0.0005) and the WM method (t_7=2.84, p=0.025).

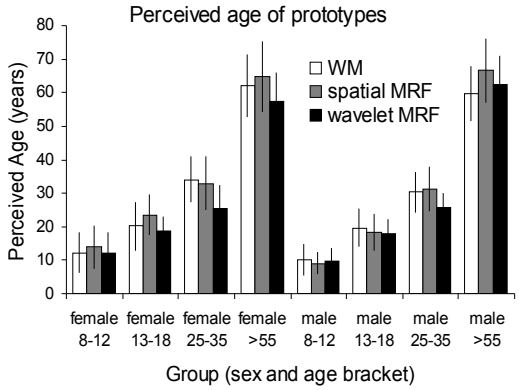

Figure 7: *Comparison of age ratings for prototypes with standard deviations.*

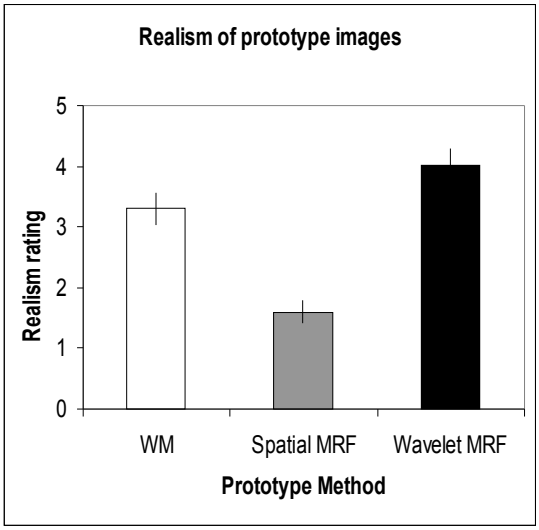

Figure 8: *Realism ratings for the wavelet magnitude, spatial MRF and wavelet MRF methods. The wavelet MRF method is significantly more realistic than the other two methods.*

4 Conclusions

We have described a new wavelet MRF based method for prototyping facial images. We have evaluated the results using a perceptual study, and the results demonstrate that the new method can produce significantly more realistic prototypes than previous methods. The new method is designed principally to aid psychology research, where the ability to capture perceived facial attributes in prototypes has allowed the design of more effective experiments into facial perception. It is hoped that this new method will add to the range and effectiveness of these psychological experiments. In addition the improved representation of texture may open up new possibilities, such as in medical imaging and cosmetics research, for capturing the typical appearance of skin conditions.

5 Acknowledgements

We would like to thank Unilever Research US for funding this research.

6 References

[PMY94] Perrett D.I., May K. and Yoshikawa S.: Attractive characteristics of female faces: preference for non-average shape. *Nature*, 368 (1994), 239-242.

[PPL*02] Perrett, D.I., Penton-Voak, I.S., Little, A.C., Tiddeman, B.P., Burt, D.M., Schmidt, N., Oxley. R., Kinloch, N. and Barrett, L.: Facial attractiveness judgements reflect learning of parental age characteristics. *Proceedings of the Royal Society, Series B.*, 269 (2002), 873-880.

[JLB*04] Jones B.C., Little A.C., Burt D.M., et al.: When facial attractiveness is only skin deep. *Perception,* 33, 5 (2004), 569-576.

[PPC*99] Penton-Voak I.S., Perrett D.I., Castles D.L., Kobayashi T., Burt D.M., Murray L.K. and Minamisawa, R.: Female preference for male faces changes cyclically. *Nature* 399: 741-742, 1999.

[DEB04] DeBruine L.M.: Resemblance to self increases the appeal of child faces to both men and women. *Evolution and Human Behaviour*, 25, 3 (May 2004), 142-154.

[RP95] Rowland D.A. and Perrett D.I.: Manipulating Facial Appearance Through Shape and Color. *IEEE Computer Graphics and Applications*, 15, 5 (1995), 70-76.

[TDR00] Tiddeman B., Duffy N. and Rabey G.: Construction and Visualisation of Three-Dimensional Facial Statistics. *Computer Methods and Programs in Biomedicine*, Vol 63 (2000), 9-20.

[TT96] Thompson P. and Toga A.W.: Visualization and mapping of anatomical abnormalities using a probabilistic brain atlas based on random fluid transformations. In Proc. VBC '96 (1996), 384-392.

[GAL78] Galton F.J.: Composite Portraits. *Nature*, Vol. 18 (1878), 97-100.

[BP93] Benson P.J. and Perrett D.I.: Extracting prototypical facial images from exemplars. *Perception*, Vol. 22 (1993), 257-262.

[TBP01] Tiddeman B., Burt D.M. and Perrett D.: Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, 21, 5 (Sept/Oct 2001), 42-50.

[MCV04] Mennucci A.C.G., Cellerino A., Valenzano D.: What is an average shape and what is an average texture? The caveats of using morphing programmes to create 'average' faces. In *Proc. HBES '04* (July 2004).

[TID04] Tiddeman B.: Blending Textured Images using a Multiscale MRF Method. In Proc. WSCG '04, 12, 3 (2004), 459-466.

[LMM95] Li H., Manjunath B.S., and Mitra S.K.: Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, vol. 57 (1995), 235–245.

[BK93] Burt P.J. and Kolczynski R.J.: Enhanced image capture through fusion. In *Proc. 4th Int. Conf. on Computer Vision*, 1993, 173–182.

[WRM95] Wilson T.A., Rogers S.K., and Myers L.R.: Perceptual based hyperspectral image fusion using multiresolution analysis. *Optical Engineering*, 34, 11 (1995), 3154–3164.

[ZB99] Zhang Z. and Blum R.: A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proceedings of the IEEE*, August 1999, 1315–1328,.

[BB04] Buntilov V. and Bretschneider T.: Assessment of the method-inherent distortions in wavelet fusion, In Proc. of the Asian Conf. on Remote Sensing, Vol. 1 (2004), 239-244.

[HCB02] Hill P., Canagarajah N. and Bull D.: Image Fusion using Complex Wavelets, *In Proc. BMVC 2002* (2002), 487-496.

[HB95] Heeger D.J. and Bergen J.R., "Pyramid-based texture analysis/synthesis," In Proc. SIGGRAPH '95 (1995), 229-238.

[SP98] Simoncelli E.P. and Portilla J.: Texture characterization via joint statistics of wavelet coefficient magnitudes. In *Proc. 5th Int. Conf. on Image Processing* (Oct 1998), Vol. 1, Chicago, IL.

[JUL62] Julesz B. , "Visual pattern discrimination," IRE Transactions on Information Theory, 1962, IT-8.

[HS81] Hassner M. and Sklansky J.: The use of Markov random fields as models of texture. In *Image Modelling*, A. Rosenfeld ed., Academic Press (New York), 1981.

[CC85a] Chellappa R. and Chatterjee S.: Texture classification using GMRF models. *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 33 (1985), 959-963.

[CC85b] Chellappa R. and Chatterjee S.: Texture synthesis using 2-D noncausal autoregressive models. *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 33 (1985), 194-203.

[CJ83] Cross G.R. and Jain A.K.: Markov random field texture models. *IEEE Trans. on PAMI*, 5, 1 (1983), 25-39.

[PL98] Paget R. and Longstaff I.D.: Texture synthesis via a noncausal nonparametric multiscale Markov random field. *IEEE Transactions on Image Processing*, Vol. 7 (1998), 925-931.

[EL99] Efros A. and Leung T.: Texture synthesis by non-parametric sampling, in Proc. 7th Int. Conf. Computer Vision, 1999, 1033-1038.

[WL00] Wei L.-Y. and Levoy M.: Fast Texture Synthesis using Tree-structured Vector Quantization. In *Proc. SIGGRAPH* '00 (2000), 479-488.

[ZWT98] Zhang J., Wang D. and Tran Q.N.: A wavelet-based multiresolution statistical model for texture. *IEEE Trans. on Image Processing*, Vol. 7 (1998) , 1621-1627.

[DEB97] De Bonet J.S.: Multiresolution sampling procedure for analysis and synthesis of texture images. In *Proceedings SIGGRAPH '97* (1997), 361-368, 1997.

[WOL90] Wolberg G., *Digital image warping*, IEEE Computer Society Press, Los Alamitos (CA), 1990.

[RM95] Ruprecht D. and Muller H.: Image warping with scattered data interpolation. *IEEE Computer Graphics and Applications*, 15, 2 (1995), 37-43.

[CTL95] Cootes T.F., Taylor C.J. and Lanatis A.: Active shape models – their training and application. *Computer Vision, Graphics and Image Understanding*, Vol. 61 (1995), 38-59.

[BP96] Beymer D. and Poggio T.: Image representation for visual learning. *Science*, Vol. 272 (1996), 1905-1909.

[VP97] Vetter T. and Poggio T.: Linear object classes and image synthesis from a single example image. *IEEE Trans. on PAMI*, Vol. 19 (1997), 733-742.

[SIL86] Silverman B.W.: *Density Estimation for Statistics and Data Analysis*, Chapman and Hall Ltd, 1986.