Joint Adaptive Colour Modelling and Skin, Hair and Clothing Segmentation Using Coherent Probabilistic Index Maps Supplementary Material

Carl Scheffler
carl.scheffler@gmail.com
Jean-Marc Odobez
odobez@idiap.ch

Idiap Research Institute
Centre du Parc
Rue Marconi 19
Martigny, Switzerland

Contents

This work was supported by the European Community's Seventh Framework Program (FP7/2007–2013) under the TA2 project (Together Anywhere, Together Anytime), grant agreement no. ICT–2007–214793, and the HUMAVIPS project (Humanoids with Auditory and Visual Abilities in Populated Spaces), grant agreement no. ICT–2009.2.1–247525.

1 Introduction

This document contains the supplementary material for the main conference article by the same title. The main purposes of this material are to show the derivations of key equations for probabilistic inference and to present more qualitative experimental results. Note that this document is not self-contained in that it references the main article in a number of places.

2 Variational Inference

Variational inference [5] is an approximate Bayesian inference scheme. The main idea is to minimise the Kullback–Leibler divergence between the full joint distribution over all latent variables in a model (which is difficult to compute) and an approximating distribution over the model variables (which is easy to compute). The approximating distribution usually factorises at least partially in the variables of the model. In this case it can be shown that updating (optimising) each factor while keeping the other factors fixed strictly improves an upper bound on the divergence between the true and approximating distributions.

In this section we derive the update equations used to infer posterior beliefs over the class of each pixel in an image, and over the discrete and continuous palettes associated with each class in an image.

2.1 Full Generative Model

In an image, I, each pixel belongs to a discrete class, z_{Ixy}, which determines the colour model used to generate the colour of the pixel, $\vec{c}_{xy}$. The classes are *skin*, *hair*, *clothing* and *background*. The x and y subscripts are the coordinates of the pixel. The class of a pixel is generated from a discrete distribution with parameters $\vec{\pi}_{xy}^{(z)}$.

There is also a Markov random field (MRF) dependence between neighbouring class pixels. The neighbours of a pixel are the 4 pixels to the north, south, east and west of it. A pixel is conditionally independent of all other pixels given its neighbours.

The model described above is referred to as a coherent probabilistic index map (CPIM) in the main article. According to the CPIM model, the class labels in each image are generated according to

$$P_{\text{CPIM}}\left(\left\langle z_{Ixy}\right\rangle \middle| \left\langle \vec{\pi}_{xy}^{(z)}\right\rangle, k_{\text{MRF}}\right) = \left[\prod_{x,y}\left[\vec{\pi}_{xy}^{(z)}\right]_{z_{Ixy}}\right] \exp\left(k_{\text{MRF}} \sum_{(xy,uv)\in\mathscr{C}} \mathbb{I}\left[z_{Ixy} = z_{Iuv}\right]\right) \tag{1}$$

$\mathscr{C}$ denotes the set of pairwise cliques — *i.e.* the set of neighbours. The square bracket notation $[\vec{\pi}]_a$ is used to refer to the value in vector $\vec{\pi}$ at index a. This is equivalent to writing π_a; the square bracket notation is used to aid legibility.

There are two types of palettes (colour models), *viz.* discrete and continuous. Each discrete palette is represented by a discrete probability vector, $\vec{\pi}_{Ik}^{(c)}$, and is generated from a Dirichlet prior with parameter $\vec{\alpha}_k$. The discrete palette for class k in image I is generated according to

$$P_{\text{pal:disc}}\left(\vec{\pi}_{Ik}^{(c)} \,\middle|\, \vec{\alpha}_k^{(c)}\right) = \text{Dirichlet}\left(\vec{\pi}_{Ik}^{(c)} \,\middle|\, \vec{\alpha}_k^{(c)}\right) \tag{2}$$

For a pixel at coordinates (x,y) that belongs to class $z_{Ixy} = k$, its colour, $\vec{c}_{Ixy}$, is drawn from the discrete probability vector $\vec{\pi}_{Ik}^{(c)}$.

$$P_{\text{col:disc}}\left(\vec{c}_{Ixy} \,\middle|\, \vec{\pi}_{Ik}^{(c)}\right) = v_{\text{bin}}^{-1}\left[\vec{\pi}_{Ik}^{(c)}\right]_{\text{bin}(\vec{c}_{Ixy})} \tag{3}$$

Here $\text{bin}(\cdot)$ is the function that maps a colour vector to its bin in the colour histogram and v_{bin} is the volume of each bin in colour space. Dividing by the volume ensures that the distribution over colour vectors is normalised. Each colour vector is a RGB triple (see Section 6 for details on colour spaces). The colour space is discretised into uniformly distributed cubes in RGB space. In this work 16 bins are used per colour dimension, for a total of $16^3 = 4096$ bins.

Each continuous palette is represented by a normal distribution over colour components in $\text{YP}_\text{b}\text{P}_\text{r}$ colour space (see Section 6). This 3-dimensional normal distribution has independent components, where each component is parametrised by its mean, μ_j, and precision (inverse variance), λ_j. The continuous palette for class k in image I is generated according to

$$\begin{aligned} & P_{\text{pal:cont}}\left(\vec{\mu}_{Ik}^{(c)}, \vec{\lambda}_{Ik}^{(c)} \,\middle|\, \vec{\eta}_k^{(c)}, \vec{\tau}_k^{(c)}, \vec{\alpha}_k^{(c)}, \vec{\beta}_k^{(c)}\right) \\ = & \prod_{j=1}^{3} \text{Normal}\left(\mu_{Ikj}^{(c)} \,\middle|\, \eta_{kj}^{(c)}, (\tau_{kj}^{(c)}\lambda_{Ikj}^{(c)})^{-1}\right) \text{Gamma}\left(\lambda_{Ikj}^{(c)} \,\middle|\, \alpha_{kj}^{(c)}, \beta_{kj}^{(c)}\right) \end{aligned} \tag{4}$$

where the gamma distribution is parametrised by its shape, α, and inverse length scale, β.[1]

For a pixel at coordinates (x,y) that belongs to class $z_{Ixy} = k$, its colour, $\vec{c}_{Ixy}$, is drawn from the normal distribution defined by $\vec{\mu}_{Ik}^{(c)}$ and $\vec{\lambda}_{Ik}^{(c)}$.

$$P_{\text{col:cont}}\left(\vec{c}_{Ixy} \,\middle|\, \vec{\mu}_{Ik}^{(c)}, \vec{\lambda}_{Ik}^{(c)}\right) = \prod_{j=1}^{3} \text{Normal}\left(c_{Ixyj} \,\middle|\, \mu_{Ikj}^{(c)}, (\lambda_{Ikj}^{(c)})^{-1}\right) \tag{5}$$

where $\vec{c}_{Ixy}$ is a vector in $\text{YP}_\text{b}\text{P}_\text{r}$ colour space.

[1] Note that the gamma distribution is often parametrised by shape, $k = \alpha$, and length scale, $\theta = \beta^{-1}$, parameters instead.

Given equations (1)–(5), the full generative model for a set of images is

$$\prod_{I\in\text{images}}\Bigg[P_{\text{CPIM}}\left(\langle z_{Ixy}\rangle \,\middle|\, \left\langle \vec{\pi}^{(z)}_{xy}\right\rangle, k_{\text{MRF}}\right)$$
$$P_{\text{pal:disc}}\left(\vec{\pi}^{(c)}_{I,\text{cloth}} \,\middle|\, \vec{\alpha}^{(c)}_{\text{cloth}}\right)$$
$$P_{\text{pal:disc}}\left(\vec{\pi}^{(c)}_{I,\text{back}} \,\middle|\, \vec{\alpha}^{(c)}_{\text{back}}\right)$$
$$P_{\text{pal:cont}}\left(\vec{\mu}^{(c)}_{I,\text{skin}}, \vec{\lambda}^{(c)}_{I,\text{skin}} \,\middle|\, \vec{\eta}^{(c)}_{\text{skin}}, \vec{\tau}^{(c)}_{\text{skin}}, \vec{\alpha}^{(c)}_{\text{skin}}, \vec{\beta}^{(c)}_{\text{skin}}\right)$$
$$P_{\text{pal:cont}}\left(\vec{\mu}^{(c)}_{I,\text{hair}}, \vec{\lambda}^{(c)}_{I,\text{hair}} \,\middle|\, \vec{\eta}^{(c)}_{\text{hair}}, \vec{\tau}^{(c)}_{\text{hair}}, \vec{\alpha}^{(c)}_{\text{hair}}, \vec{\beta}^{(c)}_{\text{hair}}\right)$$
$$\prod_{x,y}\Bigg[P_{\text{col:disc}}\left(\vec{c}_{Ixy} \,\middle|\, z_{Ixy}=\text{cloth}, \vec{\pi}^{(c)}_{I,\text{cloth}}\right)^{\mathbb{I}[z_{Ixy}=\text{cloth}]}$$
$$P_{\text{col:disc}}\left(\vec{c}_{Ixy} \,\middle|\, z_{Ixy}=\text{back}, \vec{\pi}^{(c)}_{I,\text{back}}\right)^{\mathbb{I}[z_{Ixy}=\text{back}]}$$
$$P_{\text{col:cont}}\left(\vec{c}_{Ixy} \,\middle|\, z_{Ixy}=\text{skin}, \vec{\mu}^{(c)}_{I,\text{skin}}, \vec{\lambda}^{(c)}_{I,\text{skin}}\right)^{\mathbb{I}[z_{Ixy}=\text{skin}]}$$
$$P_{\text{col:cont}}\left(\vec{c}_{Ixy} \,\middle|\, z_{Ixy}=\text{hair}, \vec{\mu}^{(c)}_{I,\text{hair}}, \vec{\lambda}^{(c)}_{I,\text{hair}}\right)^{\mathbb{I}[z_{Ixy}=\text{hair}]}\Bigg]\Bigg] \tag{6}$$

$\mathbb{I}[\cdot]$ is the indicator function, which evaluates to 1 if its argument is true and 0 if it is false. The indicator function is used to switch on exactly one of the P_{col} distributions, depending on the value of z_{Ixy}.

Note that in this model skin and hair palettes are continuous, while clothing and background palettes are discrete. For experiments where skin and hair also have discrete palettes, their factors should be adjusted appropriately in (6).

The prior hyper-parameters of this model are $\left\langle \vec{\pi}^{(z)}_{xy}\right\rangle$, k_{MRF}, $\vec{\alpha}^{(c)}_{\text{cloth}}$, $\vec{\alpha}^{(c)}_{\text{back}}$, $\vec{\eta}^{(c)}_{\text{skin}}$, $\vec{\tau}^{(c)}_{\text{skin}}$, $\vec{\alpha}^{(c)}_{\text{skin}}$, $\vec{\beta}^{(c)}_{\text{skin}}$, $\vec{\eta}^{(c)}_{\text{hair}}$, $\vec{\tau}^{(c)}_{\text{hair}}$, $\vec{\alpha}^{(c)}_{\text{hair}}$, $\vec{\beta}^{(c)}_{\text{hair}}$. Section 3 describes how their values were determined.

2.2 Approximating Distribution

The approximating distribution over the latent variables — which will be used in the variational inference framework, to produce an approximate posterior distribution — is

$$\prod_{I\in\text{images}}\left[\left[\prod_{x,y} q_{\text{class}}(z_{Ixy})\right] q_{\text{cloth}}\left(\vec{\pi}^{(c)}_{I,\text{cloth}}\right) q_{\text{back}}\left(\vec{\pi}^{(c)}_{I,\text{back}}\right)\right.$$
$$\left. q_{\text{skin}}\left(\vec{\mu}^{(c)}_{I,\text{skin}}, \vec{\lambda}^{(c)}_{I,\text{skin}}\right) q_{\text{hair}}\left(\vec{\mu}^{(c)}_{I,\text{hair}}, \vec{\lambda}^{(c)}_{I,\text{hair}}\right)\right] \tag{7}$$

with

$$\begin{aligned} q_{\text{class}}(z_{Ixy}) &= \text{Discrete}\left(z_{Ixy} \,\middle|\, \vec{\gamma}^{(z)}_{Ixy}\right) \\ &= \prod_{k\in\text{classes}}\left[\vec{\gamma}^{(z)}_{Ixy}\right]_k^{\mathbb{I}[z_{Ixy}=k]} \end{aligned} \tag{8}$$

and

$$q_k\left(\vec{\pi}^{(c)}_{Ik}\right) = \text{Dirichlet}\left(\vec{\pi}^{(c)}_{Ik} \,\middle|\, \vec{\gamma}^{(c)}_{Ik}\right) \tag{9}$$

for $k \in \{\text{cloth}, \text{back}\}$, and

$$q_k\left(\vec{\mu}^{(c)}_{Ik}, \vec{\lambda}^{(c)}_{Ik}\right) = \prod_{j=1}^{3} \text{Normal}\left(\mu^{(c)}_{Ikj} \,\middle|\, \eta'_{Ikj}, (\tau'_{Ikj}\lambda^{(c)}_{Ikj})^{-1}\right) \text{Gamma}\left(\lambda^{(c)}_{Ikj} \,\middle|\, \alpha'_{Ikj}, \beta'_{Ikj}\right) \tag{10}$$

for $k \in \{\text{skin}, \text{hair}\}$. The prime notation in the last equation is used to distinguish between the variational inference parameters and the prior hyper-parameters for continuous palettes.

2.3 Variational Update Equations

To derive the update equations for each approximating factor in (7), we compute the expectation of the log generative model under the variational beliefs over all other latent variables. See Chapter 8 of [2] for a proof of the validity of this procedure in general.

2.3.1 Update for Class Variables

In all of the equations in this sub-section, *const* refers to a term independent of k, and $\Theta^{(c)}_{Ik}$ is a placeholder for either discrete (9) or continuous (10) palette parameters. For each I, x, y and k,

$$\begin{aligned}
& \log\left[\vec{\gamma}^{(z)}_{Ixy}\right]_k \\
= \; & \text{const} + \log\left[\vec{\pi}^{(z)}_{xy}\right]_k + \int_{\Theta^{(c)}_{Ik}} q_k\left(\Theta^{(c)}_{Ik}\right) \log P_{\text{col}}\left(\vec{c}_{Ixy} \,\middle|\, \Theta^{(c)}_{Ik}\right) \\
& + \sum_{z_{I,x-1,y}} q_{\text{class}}(z_{I,x-1,y}) \sum_{z_{I,x+1,y}} q_{\text{class}}(z_{I,x+1,y}) \Bigg\{ \\
& \quad \sum_{z_{I,x,y-1}} q_{\text{class}}(z_{I,x,y-1}) \sum_{z_{I,x,y+1}} q_{\text{class}}(z_{I,x,y+1})\, k_{\text{MRF}} \sum_{(xy,uv)\in\mathscr{C}} \mathbb{I}\left[z_{Iuv} = k\right] \Bigg\} \\
= \; & \text{const} + \log\left[\vec{\pi}^{(z)}_{xy}\right]_k + \int_{\Theta^{(c)}_{Ik}} q_k\left(\Theta^{(c)}_{Ik}\right) \log P_{\text{col}}\left(\vec{c}_{Ixy} \,\middle|\, \Theta^{(c)}_{Ik}\right) + k_{\text{MRF}} \sum_{(u,v)\in\mathscr{C}_{xy}} \left[\vec{\gamma}^{(z)}_{Iuv}\right]_k
\end{aligned} \tag{11}$$

where $\mathscr{C}_{xy} = \{(x-1,y), (x+1,y), (x,y-1), (x,y+1)\}$ is the set of neighbouring locations to (x,y). For discrete palettes, $k \in \{\text{cloth}, \text{back}\}$,

$$\begin{aligned}
& \int_{\Theta^{(c)}_{Ik}} q_k\left(\Theta^{(c)}_{Ik}\right) \log P_{\text{col:disc}}\left(\vec{c}_{Ixy} \,\middle|\, \Theta^{(c)}_{Ik}\right) \\
= \; & \int_{\vec{\pi}^{(c)}_{Ik}} q_k\left(\vec{\pi}^{(c)}_{Ik}\right) \left[\log\left[\vec{\pi}^{(c)}_{Ik}\right]_{\text{bin}(\vec{c}_{Ixy})} - \log v_{\text{bin}}\right] \\
= \; & \Psi\left(\left[\vec{\gamma}^{(c)}_{Ik}\right]_{\text{bin}(\vec{c}_{Ixy})}\right) - \Psi\left(\sum_j \left[\vec{\gamma}^{(c)}_{Ik}\right]_j\right) - \log v_{\text{bin}}
\end{aligned} \tag{12}$$

For continuous palettes, $k \in \{\text{skin}, \text{hair}\}$,

$$\begin{aligned}
&\int_{\Theta_{Ik}^{(c)}} q\left(\Theta_{Ik}^{(c)}\right) \log P_{\text{col:cont}}\left(\vec{c}_{Ixy} \,\middle|\, \Theta_{Ik}^{(c)}\right) \\
= &\int_{\vec{\mu}_{Ik}^{(c)}} \int_{\vec{\lambda}_{Ik}^{(c)}} q_k\left(\vec{\mu}_{Ik}^{(c)}, \vec{\lambda}_{Ik}^{(c)}\right) \left(\sum_{j=1}^{3} \log \text{Normal}\left(c_{Ixyj} \,\middle|\, \mu_{Ikj}^{(c)}, (\lambda_{Ikj}^{(c)})^{-1}\right)\right) \\
= &-\frac{1}{2} \sum_{j=1}^{3} \int_{\mu_j} \int_{\lambda_j} \text{Normal}\left(\mu_j \,|\, \eta'_{Ikj}, (\tau'_{Ikj}\lambda_j)^{-1}\right) \\
&\qquad \text{Gamma}\left(\lambda_j \,|\, \alpha'_{Ikj}, \beta'_{Ikj}\right) \\
&\qquad \left(\log 2\pi - \log \lambda_j + \lambda_j (c_{Ixyj} - \mu_j)^2\right) \\
= &-\frac{1}{2} \sum_{j=1}^{3} \int_{\lambda_j} \text{Gamma}\left(\lambda_j \,|\, \alpha'_{Ikj}, \beta'_{Ikj}\right) \\
&\qquad \left(\log 2\pi - \log \lambda_j + \lambda_j ((\tau'_{Ikj}\lambda_j)^{-1} + (c_{Ixyj} - \eta'_{Ikj})^2)\right) \\
= &-\frac{1}{2} \sum_{j=1}^{3} \int_{\lambda_j} \text{Gamma}\left(\lambda_j \,|\, \alpha'_{Ikj}, \beta'_{Ikj}\right) \\
&\qquad \left(\log 2\pi - \log \lambda_j + {\tau'_{Ikj}}^{-1} + \lambda_j (c_{Ixyj} - \eta'_{Ikj})^2\right) \\
= &-\frac{1}{2} \sum_{j=1}^{3} \left(\log 2\pi - (\Psi(\alpha'_{Ikj}) - \log \beta'_{Ikj}) + {\tau'_{Ikj}}^{-1} + \frac{\alpha'_{Ikj}}{\beta'_{Ikj}} (c_{Ixyj} - \eta'_{Ikj})^2\right)
\end{aligned} \tag{13}$$

$\Psi(\cdot)$ is the digamma function.

2.3.2 Update for Discrete Palettes

The distributions over palettes are independent in both the generative model and the approximating distribution, when conditioned on the observed colours, $\langle \vec{c}_{Ixy} \rangle$ and the class labels, $\langle z_{Ixy} \rangle$. Because of this independence most factors in the log generative distribution (6) fold into an additive constant. For each I and each k referring to a class with a discrete palette,

$$\begin{aligned}
&\log q_k\left(\vec{\pi}_{Ik}^{(c)}\right) \\
= &\ \text{const} + \sum_{z_{I11}} \cdots \sum_{z_{IMN}} \left[\prod_{x,y} q_{\text{class}}(z_{Ixy})\right] \\
&\qquad \left[\log \text{Dirichlet}\left(\vec{\pi}_{Ik}^{(c)} \,\middle|\, \vec{\alpha}_k^{(c)}\right) + \sum_{x,y} \log \left[\vec{\pi}_{I,z_{Ixy}}^{(c)}\right]_{\text{bin}(\vec{c}_{Ixy})}\right] \\
= &\ \text{const} + \sum_{b \in \text{bins}} \left[\left[\vec{\alpha}_k^{(c)}\right]_b + \sum_{\substack{x,y: \\ \text{bin}(\vec{c}_{Ixy}) = b}} \left[\vec{\gamma}_{Ixy}^{(z)}\right]_k\right] \log \left[\vec{\pi}_{Ik}^{(c)}\right]_b
\end{aligned} \tag{14}$$

The last step follows from rearranging sums. Since (14) has the same parametric form as the log of a Dirichlet distribution, $q_k\left(\vec{\pi}_{Ik}^{(c)}\right)$ is a Dirichlet distribution with parameter

$$\left[\vec{\gamma}_{Ik}^{(c)}\right]_b = \left[\vec{\alpha}_k^{(c)}\right]_b + \sum_{\substack{x,y: \\ \mathrm{bin}(\vec{c}_{Ixy})=b}} \left[\vec{\gamma}_{Ixy}^{(z)}\right]_k \tag{15}$$

This update rule states that we add the expected counts for each colour in the class being updated to the vector of prior Dirichlet counts.

2.3.3 Update for Continuous Palettes

Since the three colour dimensions are independent, we consider them one at a time indexed by $j = 1,2,3$. For each I and each k referring to a class with a continuous palette,

$$\begin{aligned}
& \log q_k\left(\mu_{Ikj}^{(c)}, \lambda_{Ikj}^{(c)}\right) \\
= \; & \mathrm{const} + \sum_{z_{I11}} \cdots \sum_{z_{IMN}} \left[\prod_{x,y} q_{\mathrm{class}}(z_{Ixy})\right] \\
& \qquad \left[\log \mathrm{NormalGamma}\left(\mu_{Ikj}^{(c)}, \lambda_{Ikj}^{(c)} \,\middle|\, \eta_{kj}^{(c)}, \tau_{kj}^{(c)}, \alpha_{kj}^{(c)}, \beta_{kj}^{(c)}\right) + \right. \\
& \qquad \left. \sum_{x,y} \log \mathrm{Normal}\left(c_{Ixyj} \,\middle|\, \mu_{Ikj}^{(c)}, (\lambda_{Ikj}^{(c)})^{-1}\right)\right] \\
= \; & \mathrm{const} - \frac{\tau_{kj}^{(c)} \lambda_{Ikj}^{(c)}}{2}\left(\mu_{Ikj}^{(c)} - \eta_{kj}^{(c)}\right)^2 - \beta_{kj}^{(c)} \lambda_{Ikj}^{(c)} + \left(\alpha_{kj}^{(c)} - \frac{1}{2}\right) \log \lambda_{Ikj}^{(c)} \\
& + \frac{S_{Ik}^{(0)}}{2} \log \lambda_{Ikj}^{(c)} - \frac{\lambda_{Ikj}^{(c)}}{2}\left(S_{Ik}^{(0)} (\mu_{Ikj}^{(c)})^2 - 2 S_{Ikj}^{(1)} \mu_{Ikj}^{(c)} + S_{Ikj}^{(2)}\right)
\end{aligned} \tag{16}$$

where

$$\begin{aligned}
S_{Ik}^{(0)} &= \sum_{x,y} [\vec{\gamma}_{Ixy}^{(z)}]_k \\
S_{Ikj}^{(1)} &= \sum_{x,y} [\vec{\gamma}_{Ixy}^{(z)}]_k \, c_{Ixyj} \\
S_{Ikj}^{(2)} &= \sum_{x,y} [\vec{\gamma}_{Ixy}^{(z)}]_k \, c_{Ixyj}^2
\end{aligned} \tag{17}$$

Since (16) has the same parametric form as $\log q_k\left(\mu_{Ikj}^{(c)}, \lambda_{Ikj}^{(c)}\right)$, the update equations can be written down directly.

$$\begin{aligned}
\tau'_{Ikj} &= \tau_{kj} + S_{Ik}^{(0)} \\
\alpha'_{Ikj} &= \alpha_{kj} + S_{Ik}^{(0)}/2 \\
\eta'_{Ikj} &= \frac{\eta_{kj}\tau_{kj} + S_{Ikj}^{(1)}}{\tau'_{Ikj}} \\
\beta'_{Ikj} &= \beta_{kj} + \frac{1}{2}\left(S_{Ikj}^{(2)} + \eta_{kj}^2 \tau_{kj} - {\eta'_{Ikj}}^2 \tau'_{Ikj}\right)
\end{aligned} \tag{18}$$

2.3.4 Summary

Using the update equations above in an iterative scheme, marginal beliefs are computed over the following variables.

- The class of each pixel, z_{Ixy}, with posterior

$$\text{Discrete}\left(z_{Ixy} \,\middle|\, \vec{\gamma}^{(z)}_{Ixy}\right).$$

- The discrete palette (*i.e.* normalised histogram) for each discrete class, $\vec{\pi}^{(c)}_{Ik}$ with posterior

$$\text{Dirichlet}\left(\vec{\pi}^{(c)}_{Ik} \,\middle|\, \vec{\gamma}^{(c)}_{Ik}\right).$$

- The continuous palette for each continuous class, a normal distribution parameterised by $\vec{\mu}^{(c)}_{Ik}$ and $(\vec{\lambda}^{(c)}_{Ik})^{-1}$, with posterior

$$\prod_{j=1}^{3} \text{NormalGamma}\left(\mu^{(c)}_{Ikj}, \lambda^{(c)}_{Ikj} \,\middle|\, \eta'_{Ikj}, \tau'_{Ikj}, \alpha'_{Ikj}, \beta'_{Ikj}\right).$$

3 Setting the Prior Parameters

3.1 Class Variables

The discrete prior probability vector representing the belief that the pixel at coordinates (x,y) in the window around a face detection belongs to a particular class

$$\vec{\pi}^{(z)}_{xy} \tag{19}$$

is computed as the posterior when taking a Dirichlet hyper-prior of

$$\text{Dirichlet}\left(\vec{\pi}^{(z)}_{xy} \,\middle|\, C^{-1}\vec{1}\right) \tag{20}$$

where $C = 4$ is the number of classes and $\vec{1}$ is a length-C vector of ones. Computing the likelihood simply involves counting the number of times that the pixel appears in each class in the labelled training data. This gives

$$\pi^{(z)}_{xyz} = \frac{N_{xyz} + C^{-1}}{\sum_{z'} N_{xyz'} + 1} \tag{21}$$

where N_{xyz} is the number of times that the pixel at coordinate (x,y) appeared in class z in the training data.

3.2 Discrete Palettes

Referring to Figure 1 in the main article, we want to estimate the Dirichlet hyper-parameter $\vec{\alpha}^{(c)}_{k}$ based on pixels from the palette under consideration. Firstly, note that there is a latent vector for each image, $\vec{\pi}^{(c)}_{Ik}$, which can be integrated out analytically. The resulting distribution (over counts of the number of times that each discrete colour appears in an image) is known as the Pólya distribution or the Dirichlet–multinomial distribution. See [6] for details on the derivation of the Pólya distribution and the procedure for estimating $\vec{\alpha}_k^{(c)}$ from observed counts. The parameter estimation was modified slightly to deal with the large dimensionality of the colour space (it has $16^3 = 4096$ dimensions). See the `polya_distribution.py` file in the source code [1] for a Python implementation.

3.3 Continuous Palettes

To set the prior parameters over continuous colour distributions we do a maximum likelihood fit to training data. For each of the N skin colour images in the Compaq database [4], compute the sample mean, μ_{nj}, and inverse variance, λ_{nj}, of the skin pixels in the image for all 3 colour dimensions. Images are indexed by n and colour dimensions by j. Then take all sample means and inverse variances across the database as samples from the prior (4) and fit the prior parameters, $\vec{\eta}, \vec{\tau}, \vec{\alpha}, \vec{\beta}$, using maximum likelihood. By setting the derivatives of the log prior distribution to zero, the maximum likelihood setting can be found as

$$\eta_j = \frac{1}{S_j^{(\lambda)}} \sum_n \lambda_{nj} \mu_{nj} \tag{22}$$

$$\tau_j = \frac{N}{\sum_n \lambda_{nj} (\mu_{nj} - \eta_j)^2} \tag{23}$$

$$\log \alpha_j - \Psi(\alpha_j) = \log(S_j^{(\lambda)}/N) - \frac{1}{N} \sum_n \log \lambda_{nj} \tag{24}$$

$$\beta_j = N\alpha_j / S_j^{(\lambda)} \tag{25}$$

where

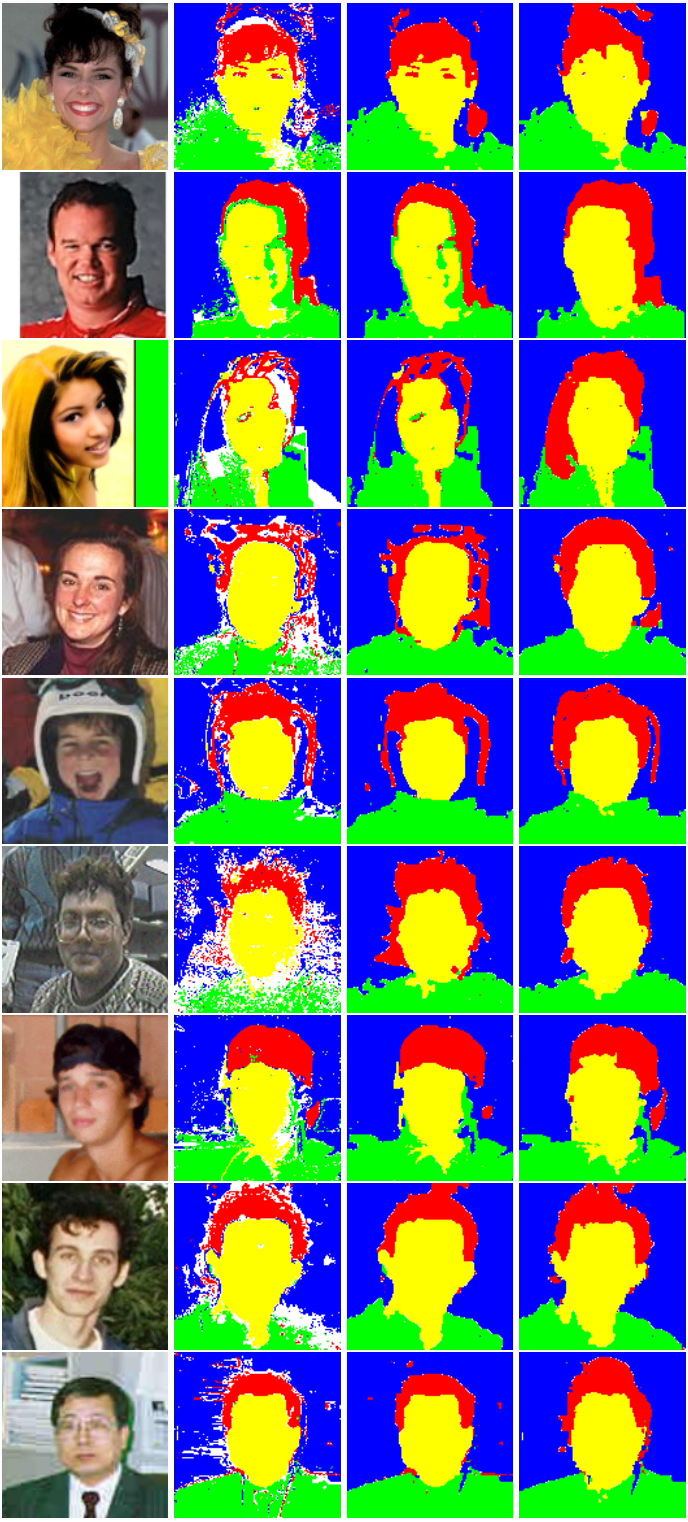
$$S_j^{(\lambda)} = \sum_n \lambda_{nj}. \tag{26}$$

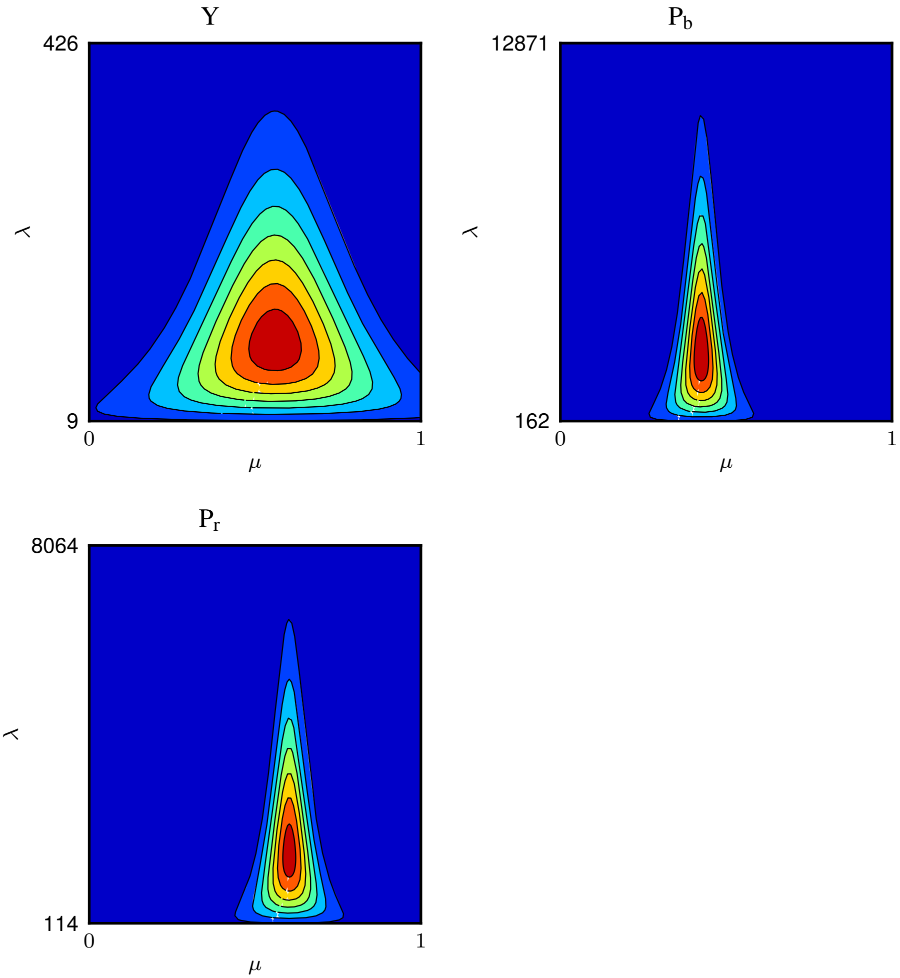
Note that (24) is an implicit equation in α_j, which is solved using a numerical root finder.

4 More Experimental Results

The figure on page 10 show some more face colour segmentation results, which were not included in the main article due to space constraints. These results are provided here without comment, to give a qualitative overview of the reliability of the model output. In each row, the first column shows the input face image (as located by the face detector [3]); the second column shows the segmentation result when the Markov random field (MRF) is off ($k_{\mathrm{MRF}} = 0$); the third column when the MRF is on ($k_{\mathrm{MRF}} = 3$) and continuous palettes are used for skin and hair; the final column when the MRF is on and discrete palettes are used for skin and hair. Red represents hair pixels, yellow skin, green clothing, and blue background. White pixels remained unclassified. Pixels were classified only if the posterior probability of the pixel belonging to a class was at least 0.75.

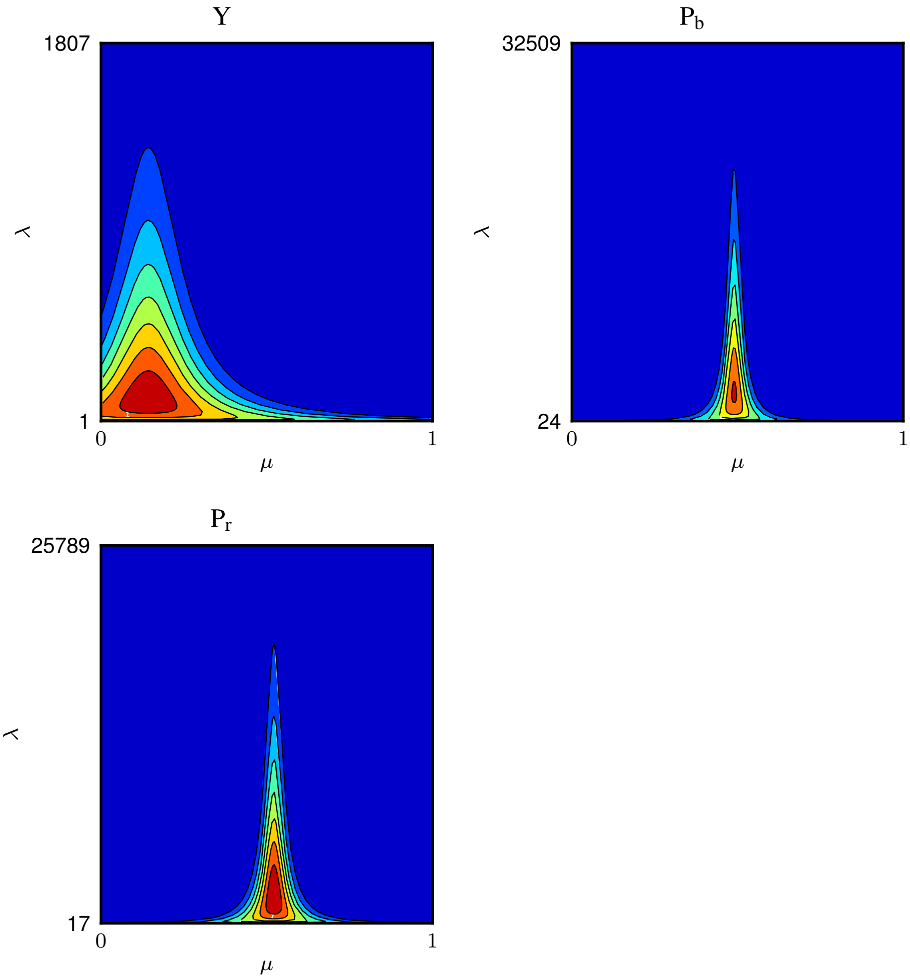
The figure on page 11 shows a visualisation of the prior belief over skin palettes. The figure on page 12 shows a visualisation of the prior belief over hair palettes. Note that the individual samples from each of these two distributions are quite distinct. Contrast this with the figure on page 13, showing samples from the posterior distributions, where the samples are all very similar.



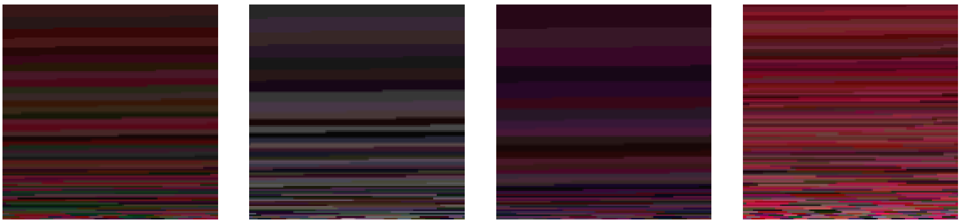


The prior belief over skin palettes. Each plot shows a probability distribution (belief) over the mean (μ) and inverse variance (λ) of the normal distribution over a colour component. The three colour components are luminance (Y), blue (P_b) and red (P_r).

Samples from the prior skin distribution. Each sample is a distribution over colours, represented here by a plot where the area devoted to each colour is proportional to its probability of being observed.



The prior belief over hair palettes. Each plot shows a probability distribution (belief) over the mean (μ) and inverse variance (λ) of the normal distribution over a colour component. The three colour components are luminance (Y), blue (P_b) and red (P_r).



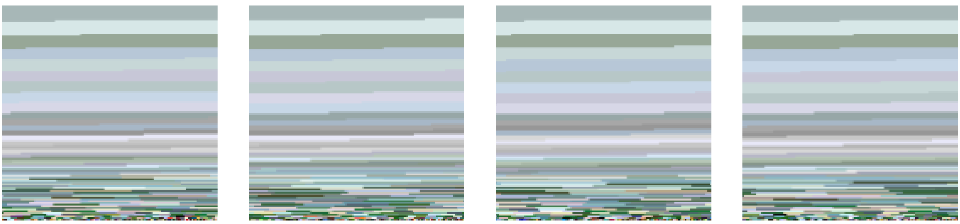
Samples from the prior hair distribution. Each sample is a distribution over colours, represented here by a plot where the area devoted to each colour is proportional to its probability of being observed.

The test image.

Samples from the posterior skin distribution.

Samples from the posterior hair distribution.

Samples from the posterior clothing distribution.



Samples from the posterior background distribution.

5 Probability Distributions

The following probability distributions are used in this article. They are stated here for completeness and to make their parametric forms explicit. Note that in the equations below, $\Gamma(\cdot)$ is the gamma function.

$$\begin{aligned}
\mathrm{Normal}\left(x \,\middle|\, \mu, \sigma^2\right) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \\
&\quad \text{with } x, \mu \in \mathbb{R} \text{ and } \sigma \in \mathbb{R}^+ && (27) \\
\mathrm{Gamma}\left(x \,|\, \alpha, \beta\right) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \text{ with } x, \alpha, \beta \in \mathbb{R}^+ && (28) \\
\mathrm{NormalGamma}\left(\mu, \lambda \,|\, \eta, \tau, \alpha, \beta\right) &= \mathrm{Normal}\left(\mu \,\middle|\, \eta, (\tau\lambda)^{-1}\right) \mathrm{Gamma}\left(\lambda \,|\, \alpha, \beta\right) && (29) \\
\mathrm{Dirichlet}\left(\vec{x} \,|\, \vec{\alpha}\right) &= \frac{\Gamma(\sum_{i=1}^d \alpha_i)}{\prod_{i=1}^d \Gamma(\alpha_i)} \prod_{i=1}^d x_i^{\alpha_i - 1} \\
&\quad \text{with } x_i, \alpha_i \in \mathbb{R}^+ \ \forall i = 1, \ldots, d \\
&\quad \text{and } \sum_{i=1}^d x_i = 1 && (30) \\
\text{Pólya}\left(\vec{n} \,|\, \vec{\alpha}\right) &= \frac{\Gamma(\sum_{i=1}^d \alpha_i)}{\Gamma\left(\sum_{i=1}^d (n_i + \alpha_i)\right)} \prod_{i=1}^d \frac{\Gamma(n_i + \alpha_i)}{\Gamma(\alpha_i)} \\
&\quad \text{with } n_i \in \mathbb{N}_0 \text{ and } \alpha_i \in \mathbb{R}^+ \ \forall i = 1, \ldots, d && (31)
\end{aligned}$$

6 Colour Spaces

6.1 RGB

This is the standard colour space used for computer displays, rescaled linearly from the range $[0, 255]$ such that

$$R, G, B \in [0, 1] \tag{32}$$

In this article, the RGB colour space is used for discrete colour models.

6.2 $\mathrm{YC_bC_r}$ / $\mathrm{YP_bP_r}$

These two colour spaces are simply scaled differently. $\mathrm{YC_bC_r}$ is used in the JPEG compression scheme.

$$\begin{aligned}
Y, C_b, C_r &\in [0, 255] && (33) \\
Y, P_b, P_r &\in [0, 1] && (34)
\end{aligned}$$

$$\begin{bmatrix} Y \\ P_b \\ P_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix} \tag{35}$$

In this article, the $\mathrm{YP_bP_r}$ colour space is used for continuous colour models.

References

[1] Adaptive face colour modelling and segmentation project page. URL http://www.idiap.ch/software/facecolormodel/.

[2] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[3] Intel Corporation. OpenCV, version 2.0.0, October 2009. URL http://opencv.willowgarage.com/wiki/.

[4] M. J. Jones and J. M. Rehg. Statistical color models with application to skin detection. *International Journal of Computer Vision*, 46:81–96, 2002.

[5] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37:183–233, 1999.

[6] T. Minka. Estimating a Dirichlet distribution. Technical report, 2009.