Robert Collins
CSE586

"Statistical Color Models with Application to Skin Detection"

M. J. Jones and J. M. Rehg

presented by Robert Collins

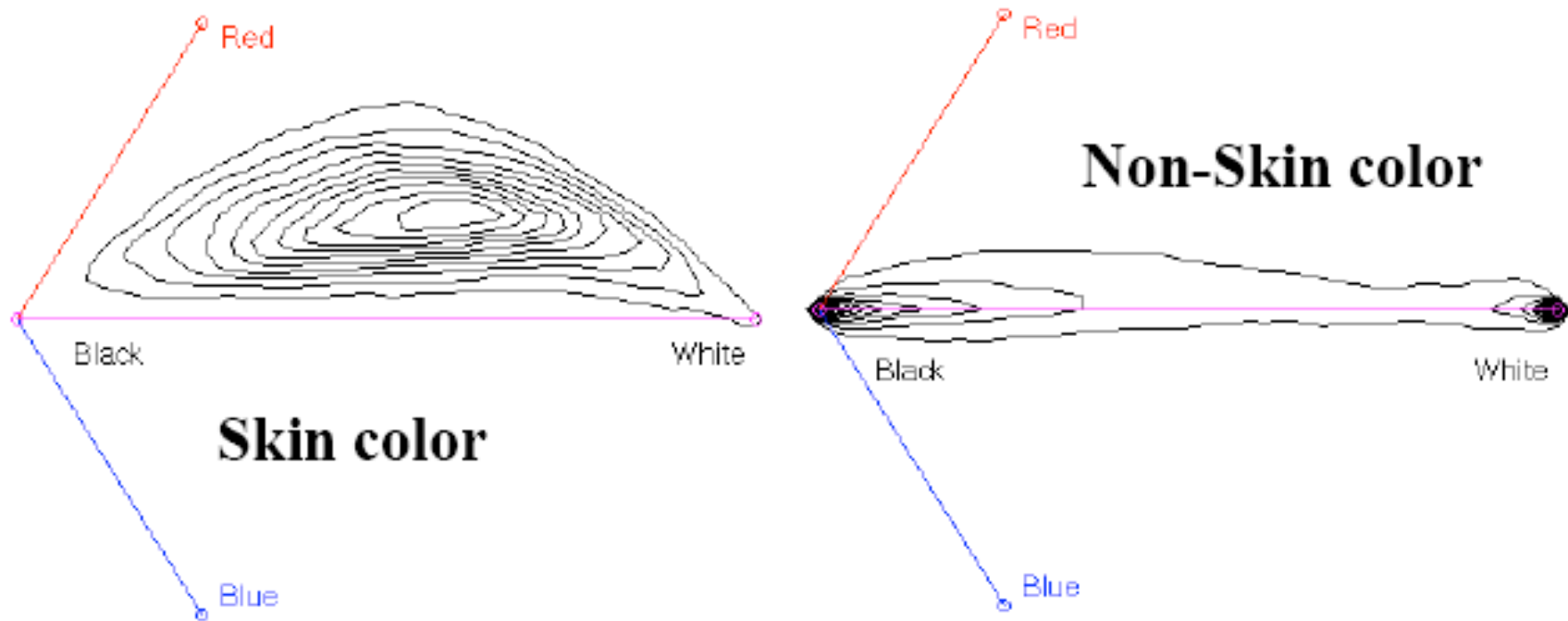# Goal: Label Skin Pixels in an Image

Applications:
Person finding/tracking
Gesture recognition
Flag possible adult content

General Overview

- **Learn distributions of skin and nonskin color**
- **Histograms; Gaussian Mixture Models (GMMs)**
- **Bayesian classification of skin pixels**
- **Combining with text-based classification**
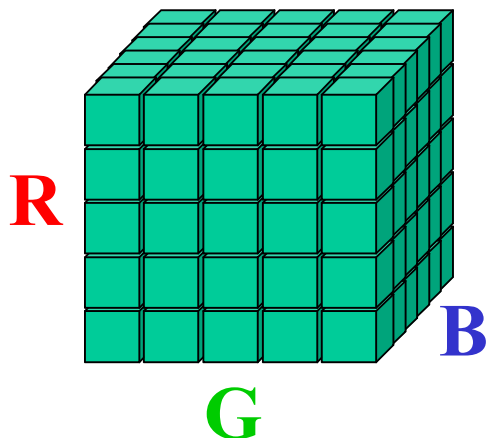
Approach:Learning from Examples

First, have some poor grad student hand label thousands of images

$$P(rgb \mid skin) = \frac{\text{number of times rgb seen for a skin pixel}}{\text{total number of skin pixels seen}}$$
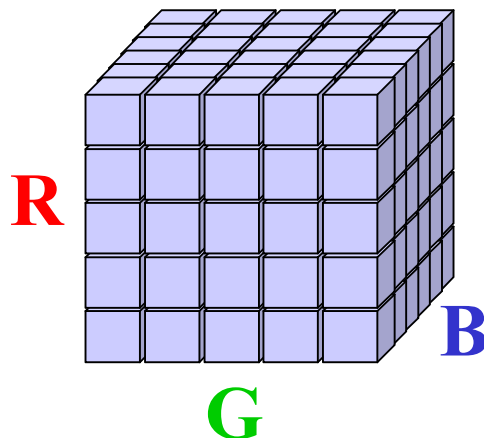
$$P(rgb \mid not\ skin) = \frac{\text{number of times rgb seen for a non-skin pixel}}{\text{total number of non-skin pixels seen}}$$

These statistics stored in two 32x32x32 RGB histograms

Skin histogram   Non-Skin histogram

Likelihood Ratio Classifier

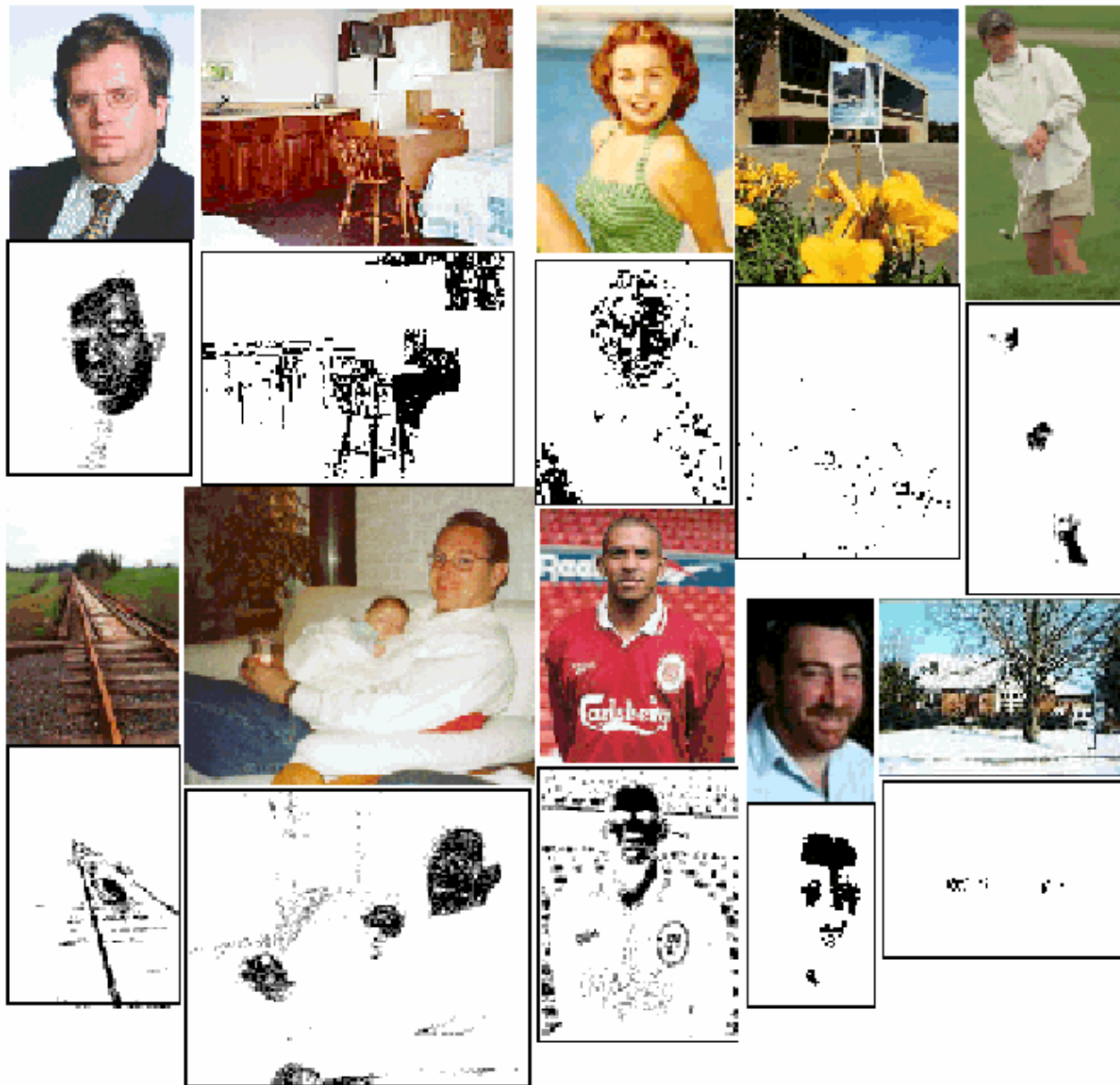**Label a pixel skin if** $\dfrac{\text{P(rgb | skin)}}{\text{P(rgb | not skin)}} > \Theta$

$$\Theta = \frac{\text{(cost of false positive) P(seeing not skin)}}{\text{(cost of false negative) P(seeing skin)}}$$

$$0 <= \Theta <= 1$$

Sample Pixel Classifications

$\Theta = .4$

Gaussian Mixture Model

A compact description is provided by converting the histogram-based model into a Gaussian Mixture model.

$$P(\mathbf{x}) = \sum_{i=1}^{N} w_i \frac{1}{(2\pi)^{\frac{3}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1}(\mathbf{x}-\mu_i)},$$

Mixture of Gaussians Skin Color Model

Mixture of Gaussians Non−skin Color Model
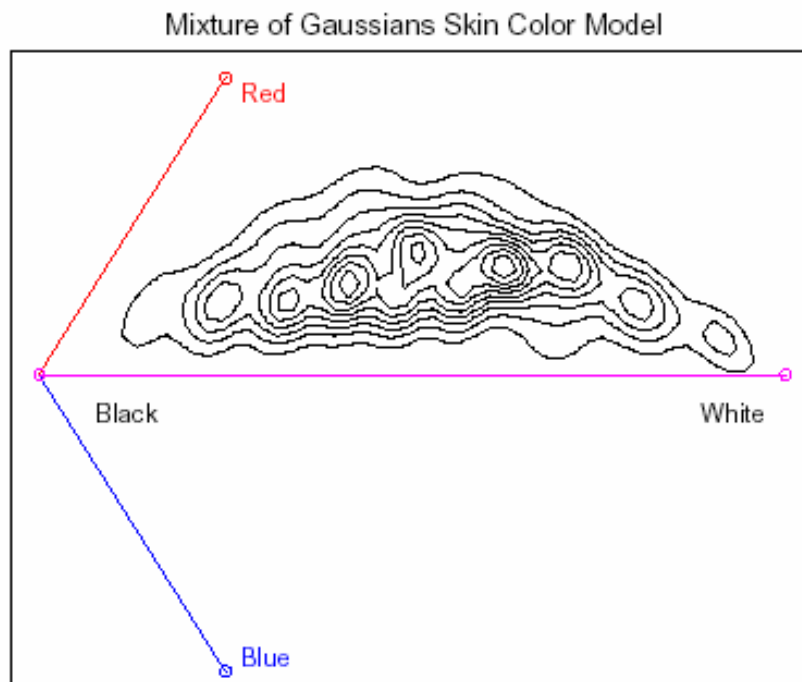
(a) Contour plot for skin model.
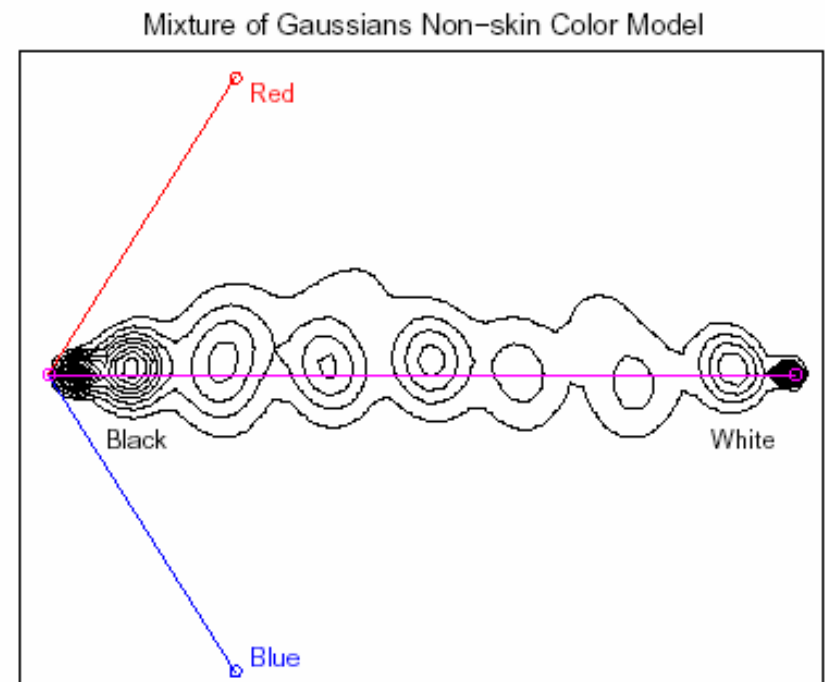
(b) Contour plot for non-skin model.

Jones and Rehg Mixture Model

Mixture of Gaussian Skin Color Model

Kernel	Mean	Covariance	Weight
1	(73.53, 29.94, 17.76)	(765.40, 121.44, 112.80)	0.0294
2	(249.71, 233.94, 217.49)	(39.94, 154.44, 396.05)	0.0331
3	(161.68, 116.25, 96.95)	(291.03, 60.48, 162.85)	0.0654
4	(186.07, 136.62, 114.40)	(274.95, 64.60, 198.27)	0.0756
5	(189.26, 98.37, 51.18)	(633.18, 222.40, 250.69)	0.0554
6	(247.00, 152.20, 90.84)	(65.23, 691.53, 609.92)	0.0314
7	(150.10, 72.66, 37.76)	(408.63, 200.77, 257.57)	0.0454
8	(206.85, 171.09, 156.34)	(530.08, 155.08, 572.79)	0.0469
9	(212.78, 152.82, 120.04)	(160.57, 84.52, 243.90)	0.0956
10	(234.87, 175.43, 138.94)	(163.80, 121.57, 279.22)	0.0763
11	(151.19, 97.74, 74.59)	(425.40, 73.56, 175.11)	0.1100
12	(120.52, 77.55, 59.82)	(330.45, 70.34, 151.82)	0.0676
13	(192.20, 119.62, 82.32)	(152.76, 92.14, 259.15)	0.0755
14	(214.29, 136.08, 87.24)	(204.90, 140.17, 270.19)	0.0500
15	(99.57, 54.33, 38.06)	(448.13, 90.18, 151.29)	0.0667
16	(238.88, 203.08, 176.91)	(178.38, 156.27, 404.99)	0.0749

Jones and Rehg Mixture Model

Mixture of Gaussian Non-skin Color Model

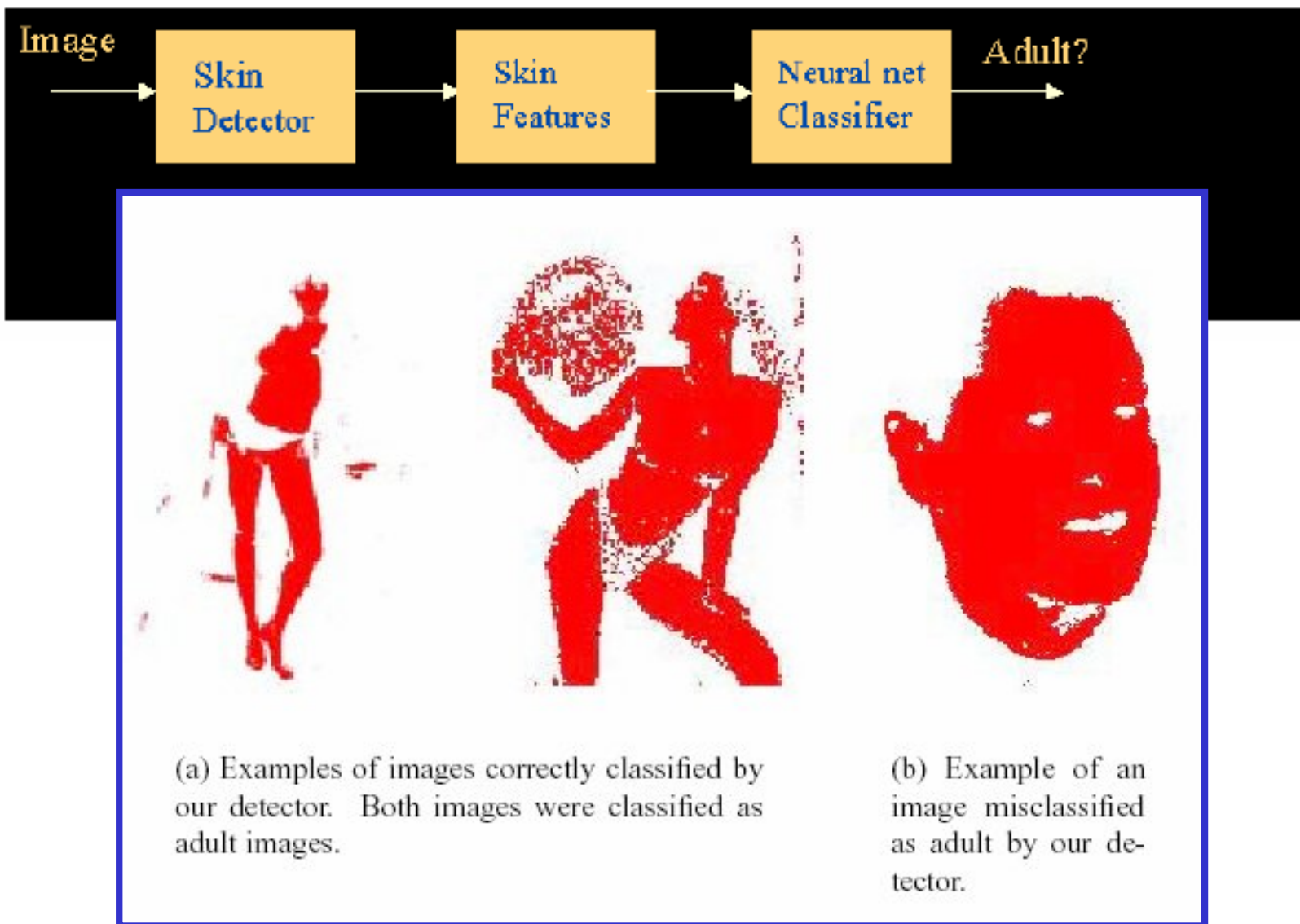| Kernel	Mean	Covariance	Weight
1	(254.37, 254.41, 253.82)	(2.77, 2.81, 5.46)	0.0637
2	(9.39, 8.09, 8.52)	(46.84, 33.59, 32.48)	0.0516
3	(96.57, 96.95, 91.53)	(280.69, 156.79, 436.58)	0.0864
4	(160.44, 162.49, 159.06)	(355.98, 115.89, 591.24)	0.0636
5	(74.98, 63.23, 46.33)	(414.84, 245.95, 361.27)	0.0747
6	(121.83, 60.88, 18.31)	(2502.24, 1383.53, 237.18)	0.0365
7	(202.18, 154.88, 91.04)	(957.42, 1766.94, 1582.52)	0.0349
8	(193.06, 201.93, 206.55)	(562.88, 190.23, 447.28)	0.0649
9	(51.88, 57.14, 61.55)	(344.11, 191.77, 433.40)	0.0656
10	(30.88, 26.84, 25.32)	(222.07, 118.65, 182.41)	0.1189
11	(44.97, 85.96, 131.95)	(651.32, 840.52, 963.67)	0.0362
12	(236.02, 236.27, 230.70)	(225.03, 117.29, 331.95)	0.0849
13	(207.86, 191.20, 164.12)	(494.04, 237.69, 533.52)	0.0368
14	(99.83, 148.11, 188.17)	(955.88, 654.95, 916.70)	0.0389
15	(135.06, 131.92, 123.10)	(350.35, 130.30, 388.43)	0.0943
16	(135.96, 103.89, 66.88)	(806.44, 642.20, 350.36)	0.0477

Sample Use: Adult Image Classification
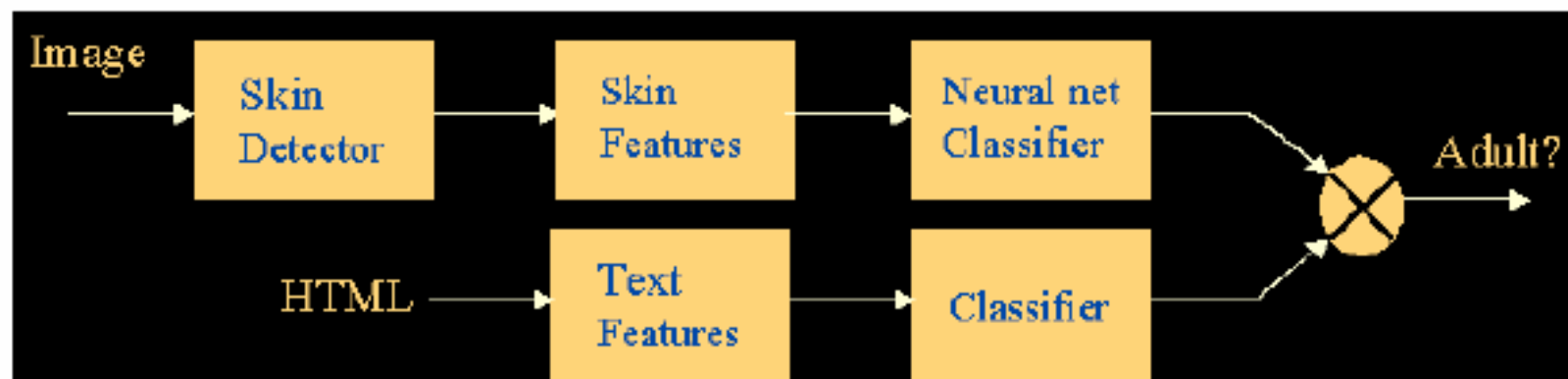

Based on Five Features:

- Percentage of pixels detected as skin.
- Average probability of the skin pixels.
- Size in pixels of the largest connected
    component of skin.
- Number of connected components of skin.
- Percentage of colors with no entries in the
    skin and non-skin histograms

Adult Image Classification

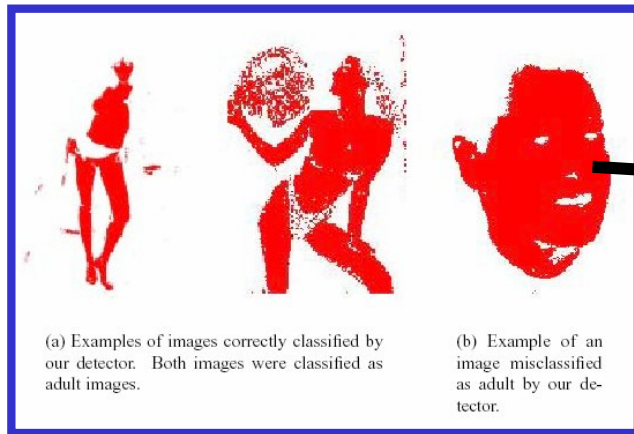(a) Examples of images correctly classified by our detector. Both images were classified as adult images.

(b) Example of an image misclassified as adult by our detector.

Combining Color and Text

	% correctly detected adult images	% false alarms
Color-based Detector	85.8%	7.5%
Text-based Detector	84.9%	1.1%
Combined Detector	93.9%	8.0%

Lessons Learned

- Harness the web as a source of data!

- With enough data, even simple learning methods based on counting can produce good classification results

- Likelihood ratio is important – model both the object AND not-object distributions to avoid thresholds on raw probabilities.

- EM and GMM models used to encode compact descriptions of color histograms.

Questions

(a) Examples of images correctly classified by our detector. Both images were classified as adult images.

(b) Example of an image misclassified as adult by our detector.

Can we combine with a face detector so portraits do not cause false positives?

I wonder how k-nearest-neighbor classification would work?