## 1. Motivation

Use correlations in group shots to help solve unconstrained face recognition.

- **Buffy** is darker-skinned than **Willow**.
  - This is detectable despite drastic changes in lighting, pose and expression.

- **Giles** is taller than **Buffy**.
  - This holds true across widely varying viewing angles, pose and location.

## 2. Features

### Attributes

- Many attributes are highly correlated with each other.

### Color-based descriptor

- Median HSV color is extracted around fiducial points: left and right eyes, mouth and convex hull of the fiducials.

### Height-based descriptor

- Height is computed as the ratio of the y-location of the face to the height of the face box.

## 3. Baseline Model

### One Gaussian per individual

$$P(Buffy \mid TestFeature) = N(\mu_{Buffy}, \Sigma_{Buffy})$$

- Each individual’s Gaussian model is characterized by a mean $$\mu$$ and a full covariance matrix $$\Sigma$$.

- **Note:** Assumes all features are independent between people.

## 4. Joint Model

### One Gaussian per pair of individuals

$$P(Buffy, Riley \mid TestFeatures) = 1/2 N(\mu_{Buffy,Riley}, \Sigma_{Buffy,Riley})$$

- $$\Sigma$$ is a diagonal co-variance matrix due to data scarcity.

- **Examples:**
  - Joyce smiles when her daughter **Dawn** is present.
  - Buffy appears darker and blurry when Spike (a vampire) is present (usually a night-time fight scene).

## 5. Conditional Model

### One Gaussian & one probability vector per pair of individuals

- **Examples:**
  - Giles is taller than Buffy.
  - Buffy is darker-skinned than Willow.
  - Dawn has a smaller nose than Riley.

- Can transitively infer relationships: Giles & Dawn not seen together, but
  - Height(Giles) > Height(Buffy) & Height(Buffy) > Height(Dawn).
  - So, Height(Giles) > Height(Dawn).

With ANY existing system in place of the baseline model, the conditional model can provide an accuracy boost for group shots.

Features for conditional and baseline models can be different.

## 6. Datasets

- **Buffy the Vampire Slayer**
  - *N. Ewing, J. Hsu, A. Belhumeur, A. Zisserman. Buffy – A Personal Photo Album. BMVC, 2006.*

- **A Personal Photo Album**
  - *~6000 frames, 2 episodes.*

## 7. Results

### Accuracy on color and height descriptors

![Accuracy on color and height descriptors](image)

- **Joint model does much better than baseline.**

### Accuracy vs. number of attributes

![Accuracy vs. number of attributes](image)

- **Joint model does generally better than baseline.**

- **Conditional model always does better than baseline.**

- **Gains increase with more attributes.**

## 8. Conclusions

- Individually, attributes outperform weaker color and height based descriptors.
- Color and height are usually uninformative for a single face, but can help recognition when considered in a joint setting.
- The Joint model suffers due to data scarcity and is outperformed by the versatile Conditional model.

---

Acknowledgment: This work was supported in part by ONR MURI Grant N00014-08-1-0638.