Ranking and Classifying Attractiveness of Photos in Folksonomies

Jose San Pedro Wandelmer (jsanpedro@mac.com)
The University of Sheffield, UK

Stefan Siersdorfer (siersdorfer@l3s.de)
L3S Research Center, Hannover, Germany

Motivation

Web 2.0 is hot!

- Easy
- Collaborative
- Dynamic

Massive amount of shared resources

- pose problems for effective search & retrieval
- especially critical for multimedia information
  - 'Semantic gap' limits the effectiveness of content-based techniques
- Finding relevant content becomes a difficult task

Motivation

Vast amounts of user generated data available

- Metadata
  - tags, descriptions, comments, etc.
- User feedback
  - implicit: n. views, n. comments, etc.
  - explicit: ratings, favorite assignments, etc.
- Not always reliable: irregularities, sparsity, etc

Can we combine user generated data and content features to enhance retrieval?

- 'Community knowledge' can help to learn about the content
- Content can help overcome irregularities of user data

Problem Setting

Scenario

- Photo sharing - Flickr

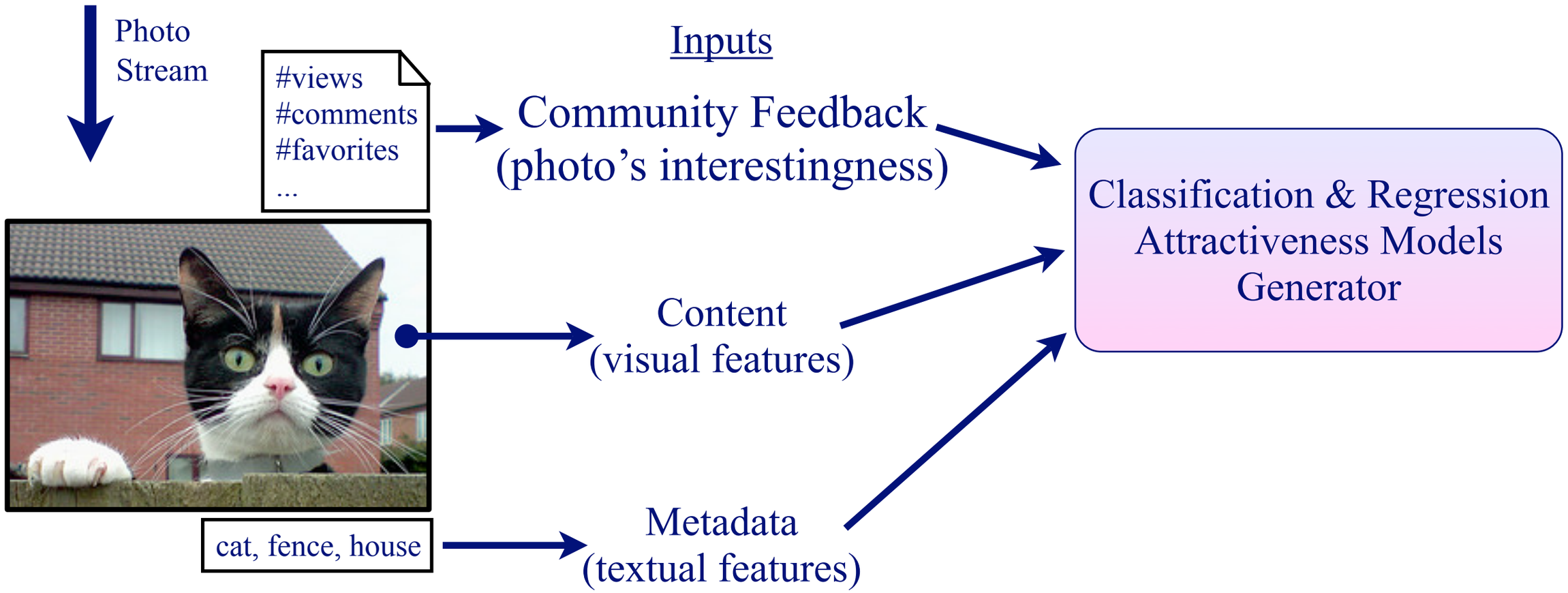
Objective

- Determine *attractiveness* of shared photos
- Reasons
  - Direct application for photo retrieval enhancement
  - Subjective concept
    - but there is a whole community providing judgements
  - Image semantics are not critical
    - enables efficient use of content-based visual features

Problem Setting

We propose

- A methodology for classification and ranking of images based on their visual appeal



Attractiveness of Images

Factors that influence human perception of attractiveness?

Landscape

Portrait

Flower

Attractiveness Visual Features

Human visual perception mainly influenced by

- Color distribution
- Coarseness

These are complex concepts

- Convey multiple orthogonal aspects
- Necessity to consider different low level features

Attractiveness Visual Features

Color Features

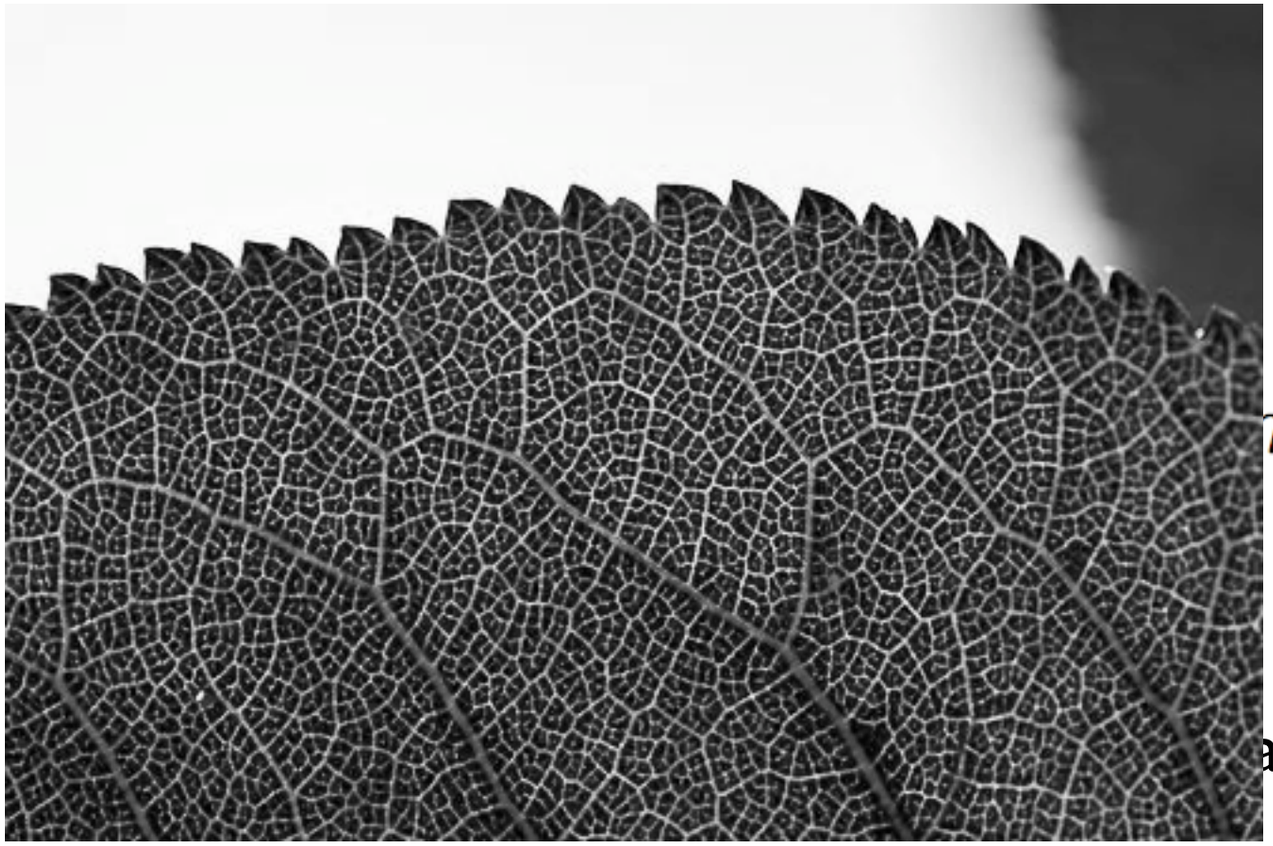
- Brightness
- Contrast
  - Luminance, RGB
- Colorfulness
- Naturalness

- Saturation $S = max(R, G, B) - min(R, G, B)$
  - Mean, Variance
  - Intensity of the colors
  - Saturation is 0 for grey scale images

Attractiveness Visual Features

Color Features

- Brightness
- Contrast
  - Luminance, RGB



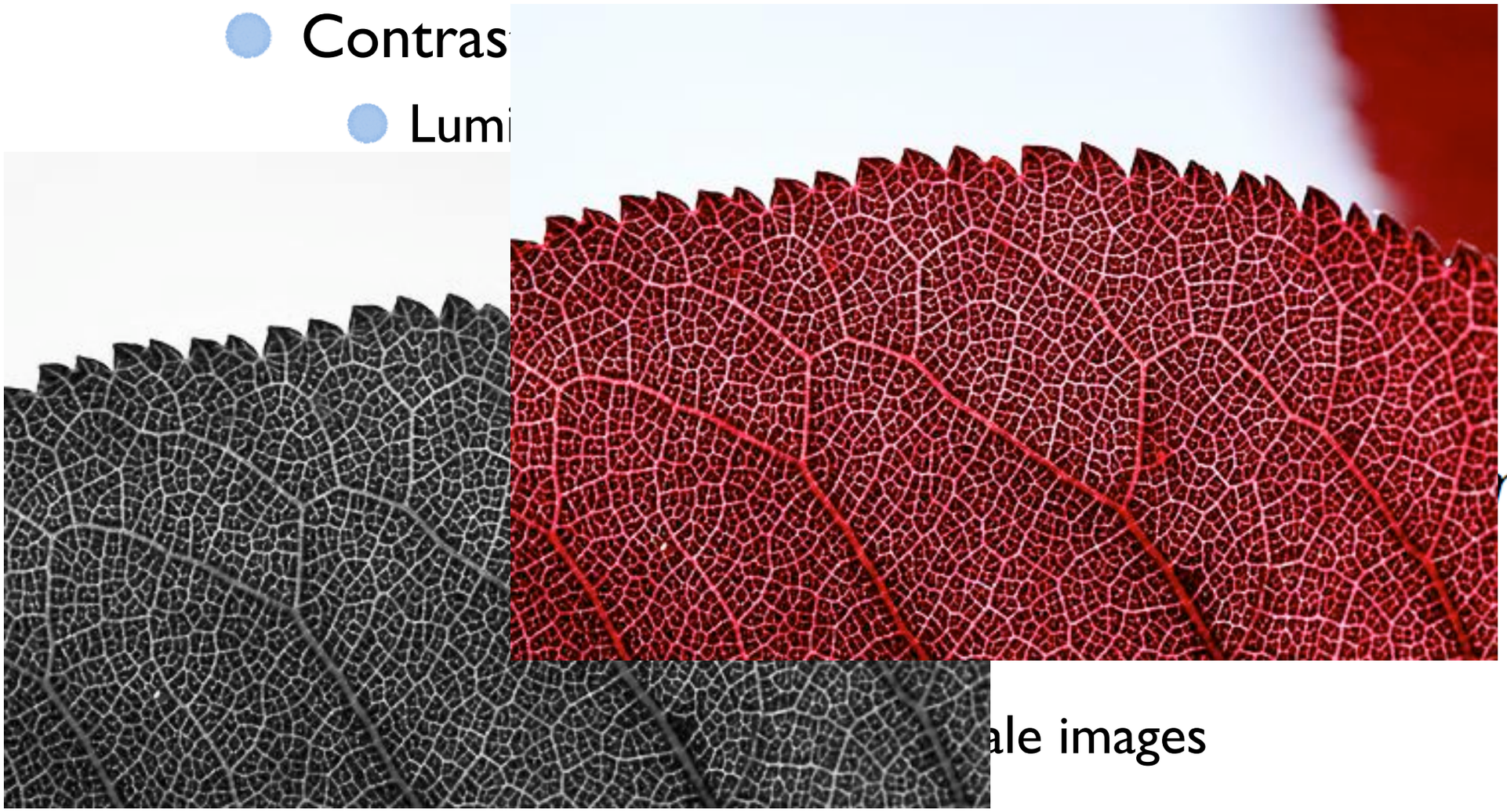
$$nax(R, G, B) - min(R, G, B)$$

ale images

Attractiveness Visual Features

Color Features

- Brightness
- Contras
  - Lumi



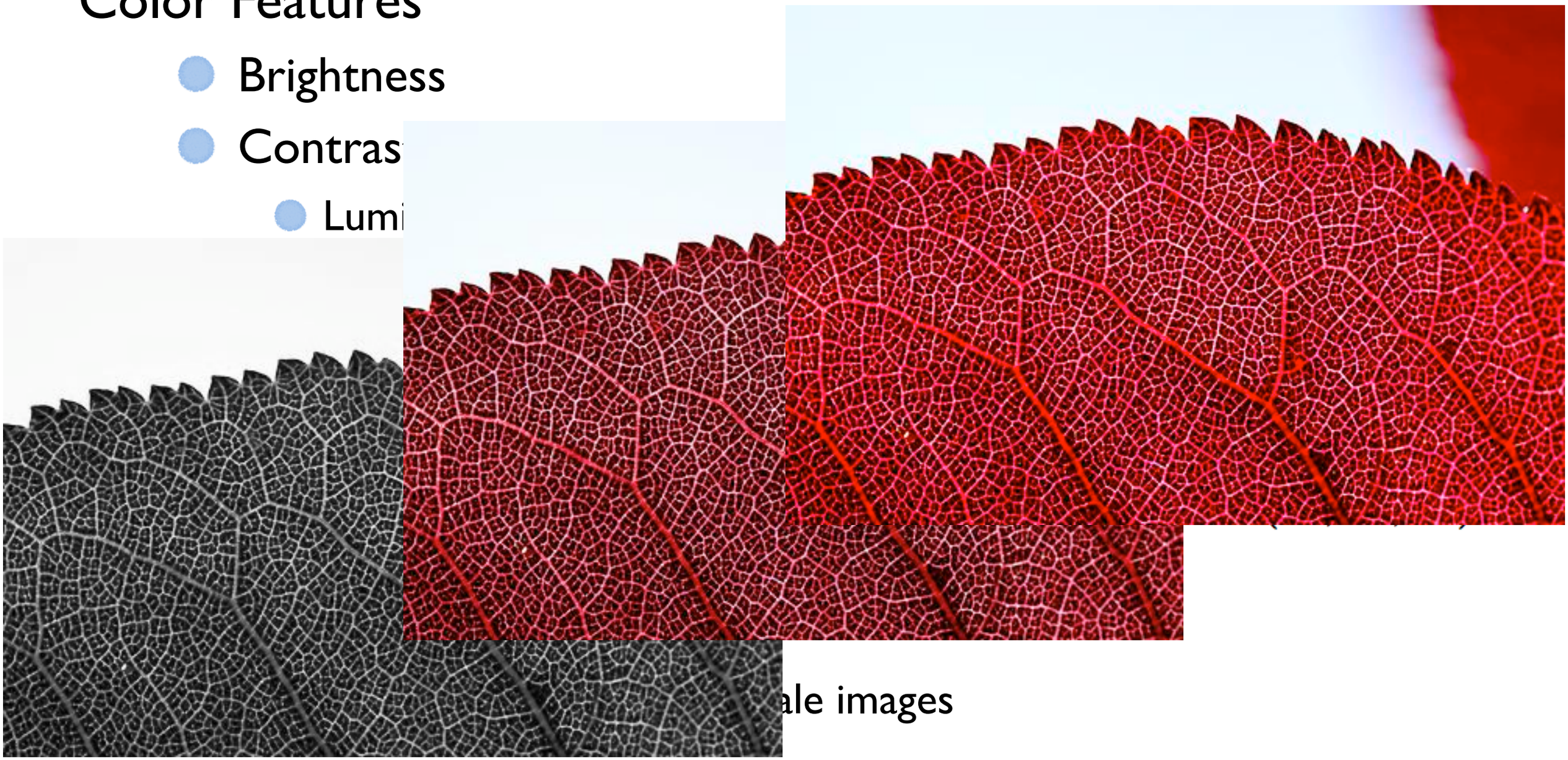
$nin(R, G, B)$

ale images

Attractiveness Visual Features

Color Features

- Brightness
- Contras
  - Lumi



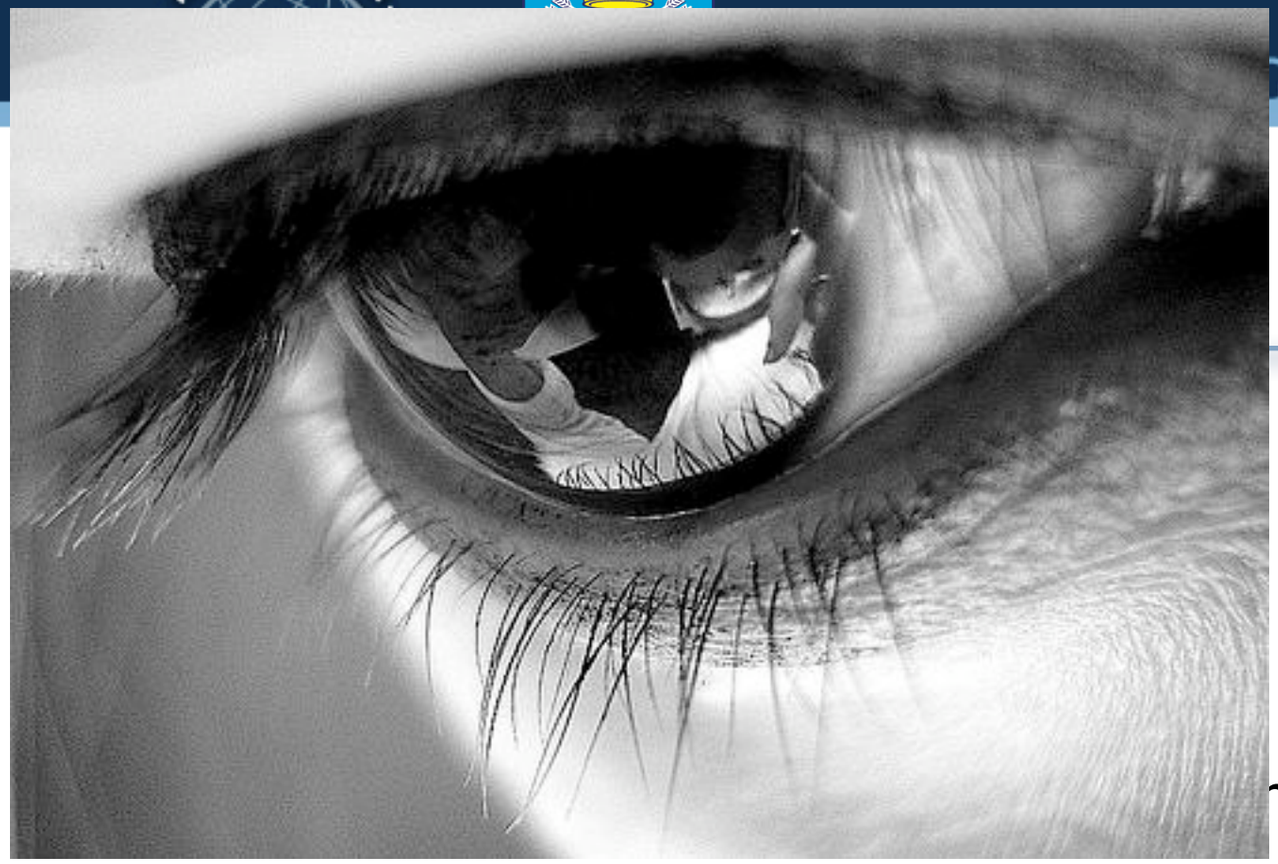
ale images

Visual Features

Coarseness

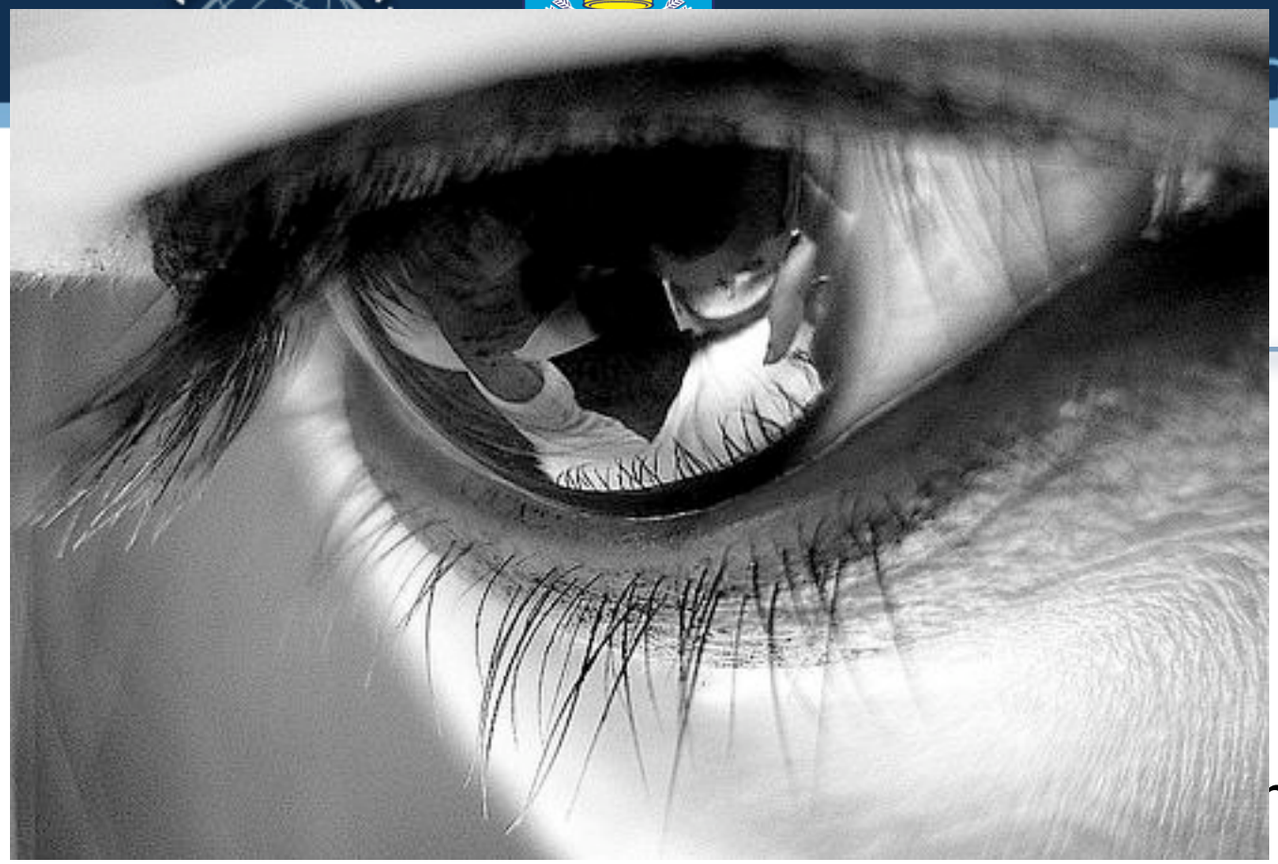
- Resolution + Acutance
- Sharpness
  - Critical importance for final appearance of photos [Savakis 2000]

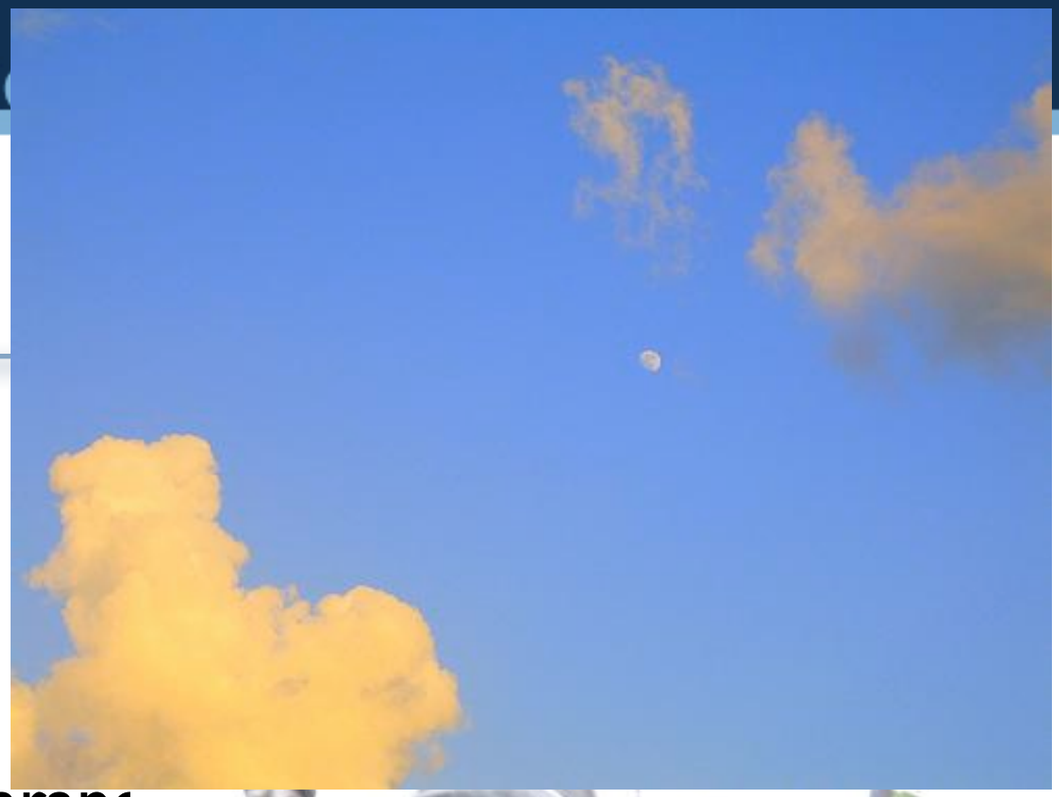
$$Sh = \sum_{x,y} \frac{L(x,y)}{\mu_{xy}}, \text{ with } L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$



nal appearance of photos [Savakis 2000]

h $L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$





nal appearanc

h $L(x, y)$

Textual Features

We consider user generated meta data

- Correlation of topics with image appealing
- Tags seem appropriate to capture this information

	technical	***judgements***	***awards***	***concepts***
Attractive Pictures	macro portrait canon landscape hdr	coolest bravo art supershot superbmasterpiece	diamondclassphotographer flickrdiamond blueribbonwinner colorphotoaward	nature sky sunset woman sea

	dates	***events***	***places***	***people***
Non Attractive Pictures	2007 2006 may	wedding graduation honeymoon birthday	madagascar prague china vegas	matt ian dad

Classification & Regression Models

Classification

- Automatic categorization of photos
  - attractive vs unattractive
- Supervised learning paradigm
  - Training set of photos represented as feature vectors
    - numeric values of visual features
    - tfidf weights of tags
  - We consider different vector representations
    - Text features only - Visual features only - Text+Visual features
  - Necessity of a sufficiently large training set of labeled photos
    - Flickr provides large photo collections with social feedback
    - Number of favorites (*NumFav*) as interestingness indicator
      - Distinct thresholds for minimum NumFav

Classification & Regression Models

Classification

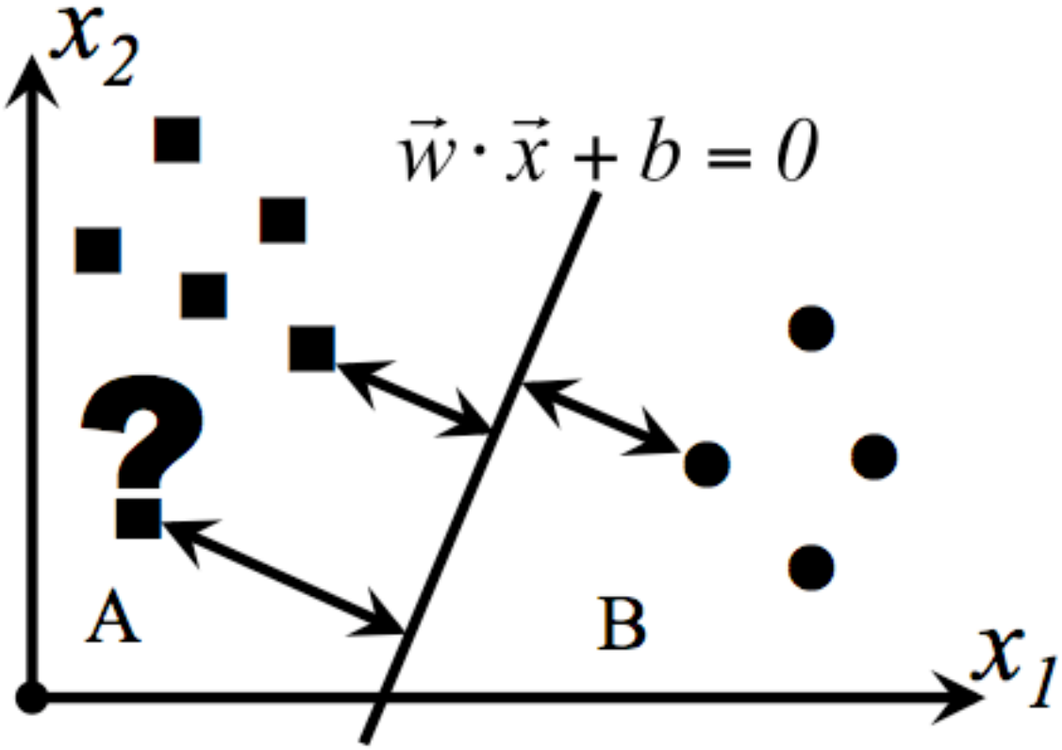
- Formally: $\{(\vec{p_1}, l_1), \ldots, (\vec{p_n}, l_n)\}$ $l_i = \begin{cases} 1, & \text{NumFav} > \text{thr} \\ -1, & \text{Otherwise} \end{cases}$

- Linear SVMs
  - Training:

$$w \cdot x + b = 0$$

  - Classification:

$$\vec{w} \cdot \vec{y} + b > 0\ ?$$



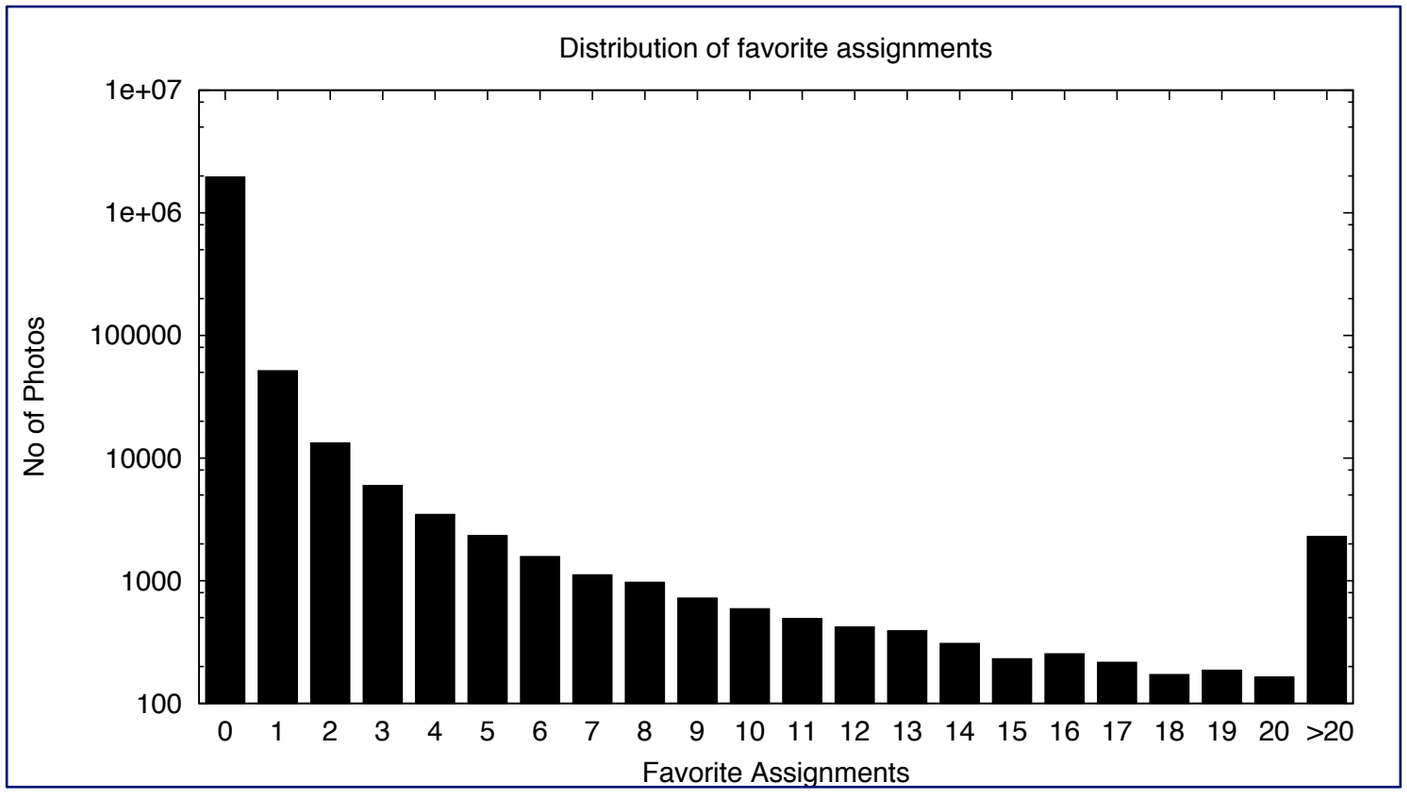
Regression

- Formally: $\{(\vec{p_1}, r_1), \ldots, (\vec{p_n}, r_n)\}$
- SVM Regression
  - Find a function to assign continuous relevance values

Experiments

Data

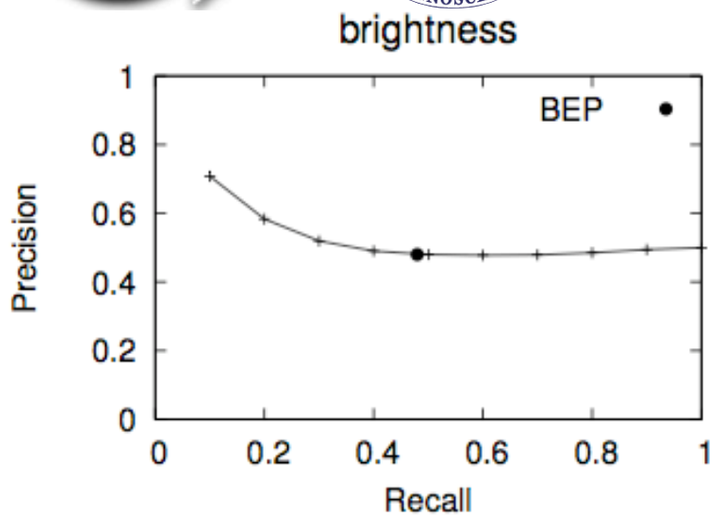
- Sample of Flickr photos
- Uploaded between June 1st and 7th 2007
- Flickr API for new photos - 20 minutes time interval
- Collection size: 2.2M - N. Users: 185k
- Test sets
  - Positive examples
    - N. Fav >= 2
    - Size : 35,000
  - Negative examples
    - N. Fav = 0
    - Size : 40,000

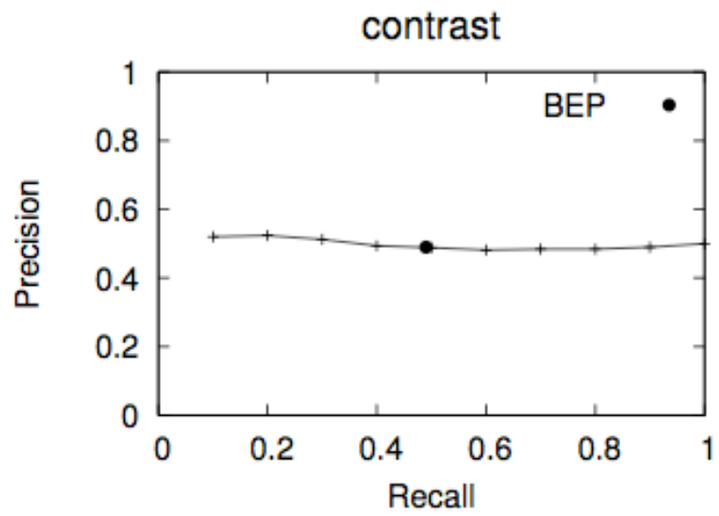


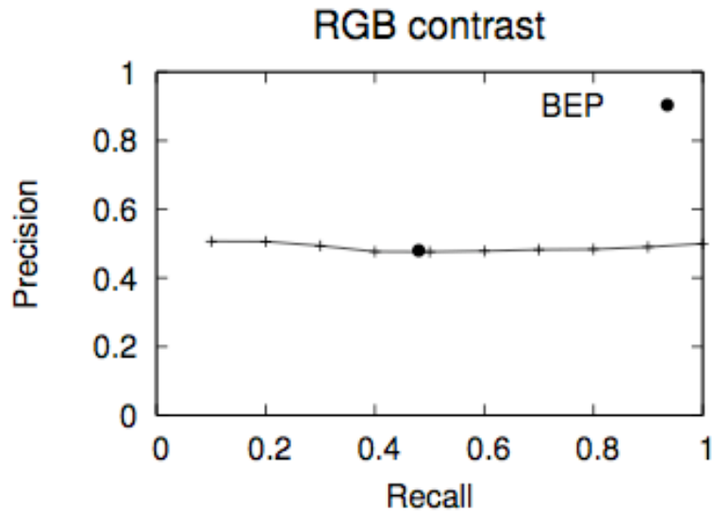
Experiments

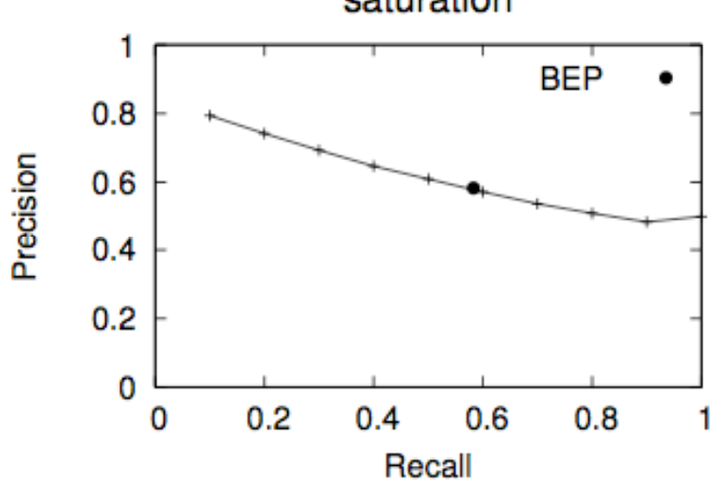
Classification

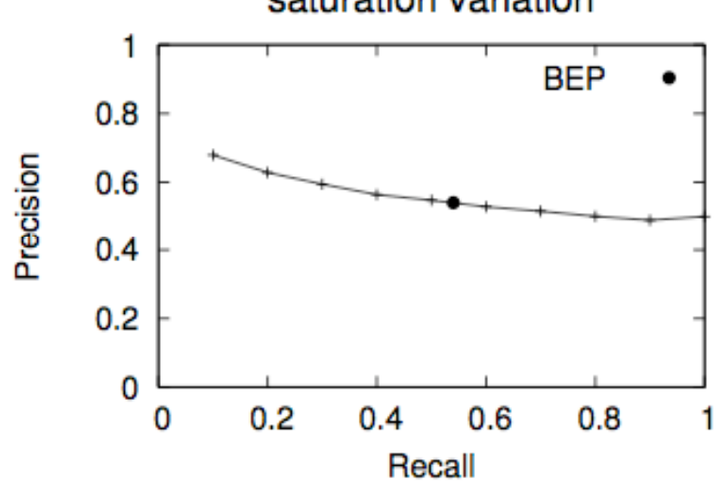
- Setup
  - Attractive defined at various restrictiveness levels:
    - NumFav >= 2, 5, 10, and 20
  - Random selection of training photos:
    - Set size = 500, 2000, 8000, and 20000
- Results
  - T=8000, NumFav >= 5
    - BEP for different configurations
      - Text Features : 0.7843
      - Visual Features: 0.6664
      - Text + Visual : 0.8363

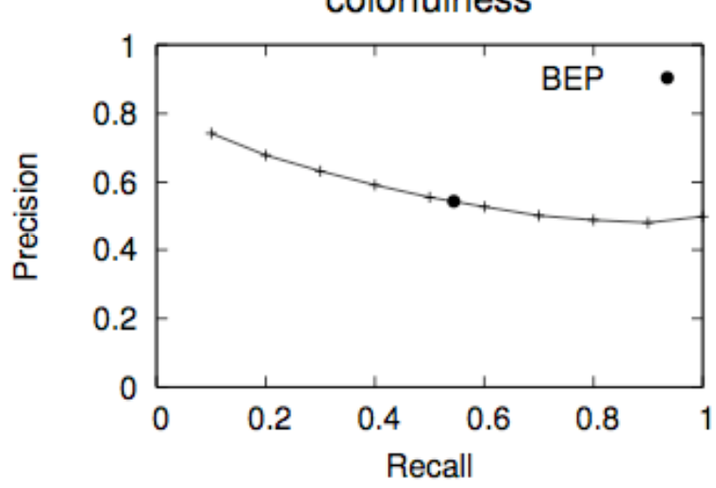

brightness
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1

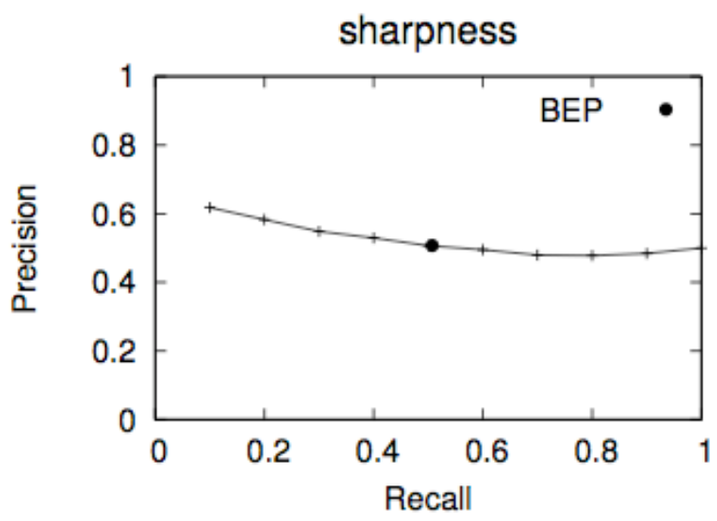

contrast
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1

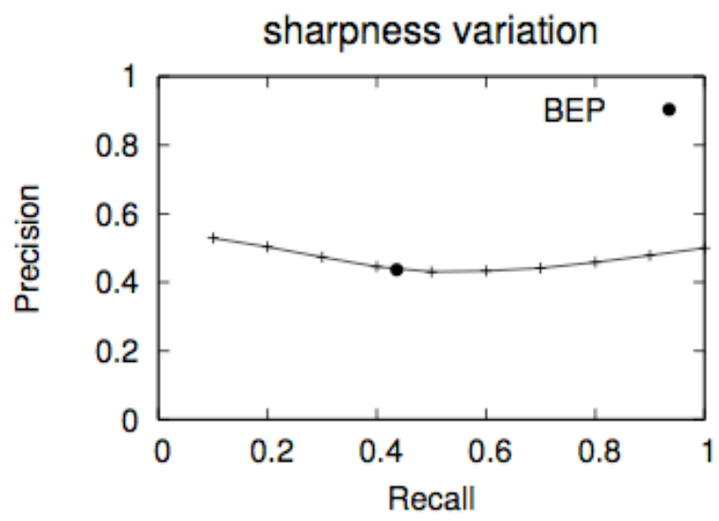

RGB contrast
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1

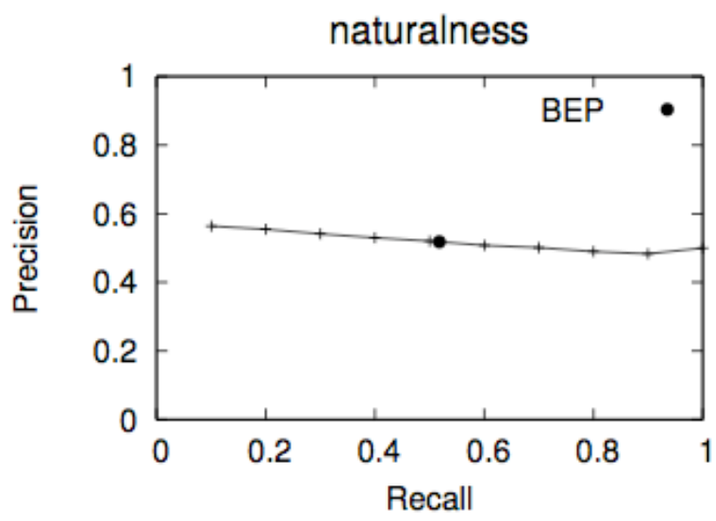

saturation
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1


saturation variation
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1


colorfulness
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1


sharpness
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1


sharpness variation
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1

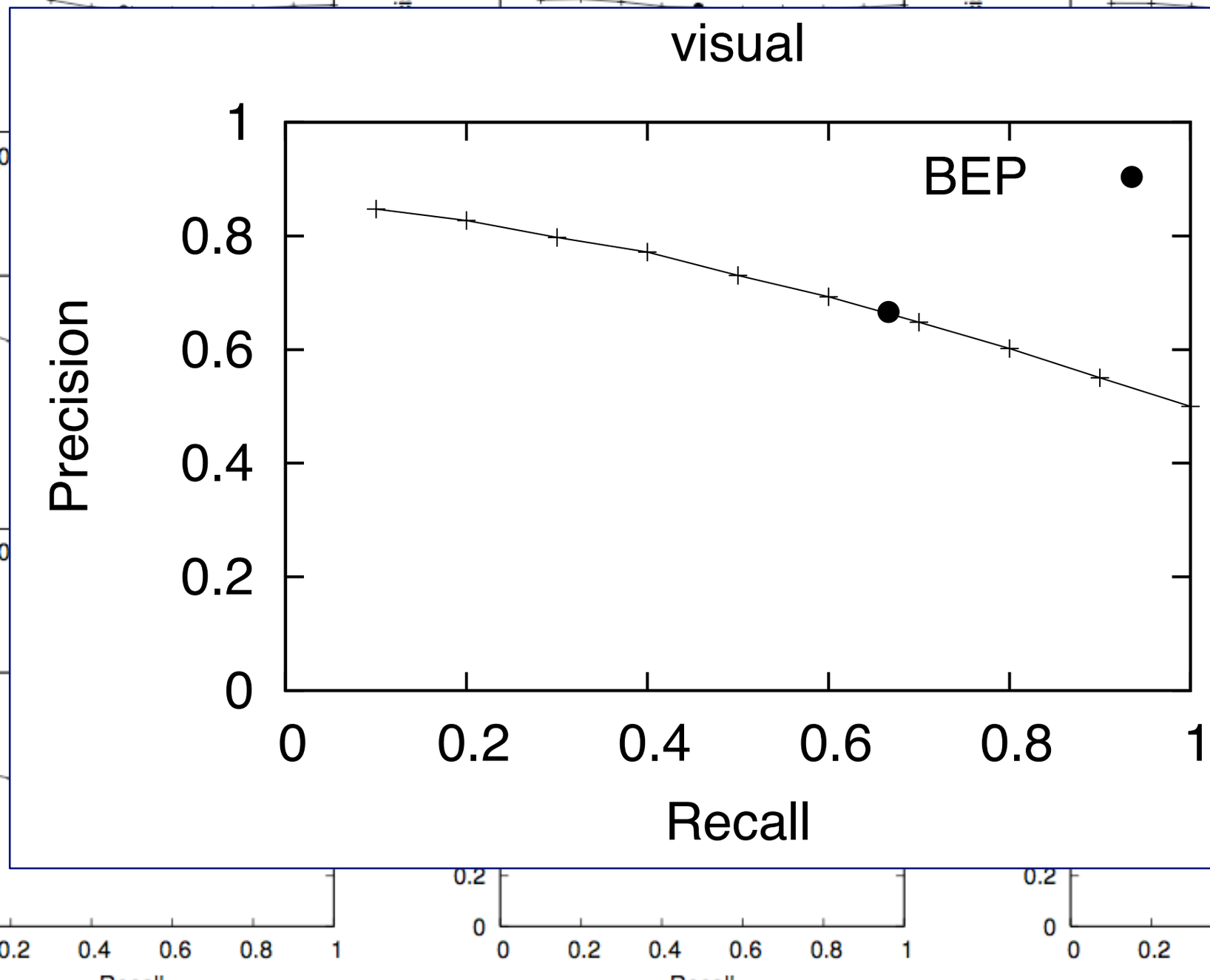

naturalness
BEP
Precision
Recall
0
0.2
0.4
0.6
0.8
1

brightness

contrast

RGB contrast

visual

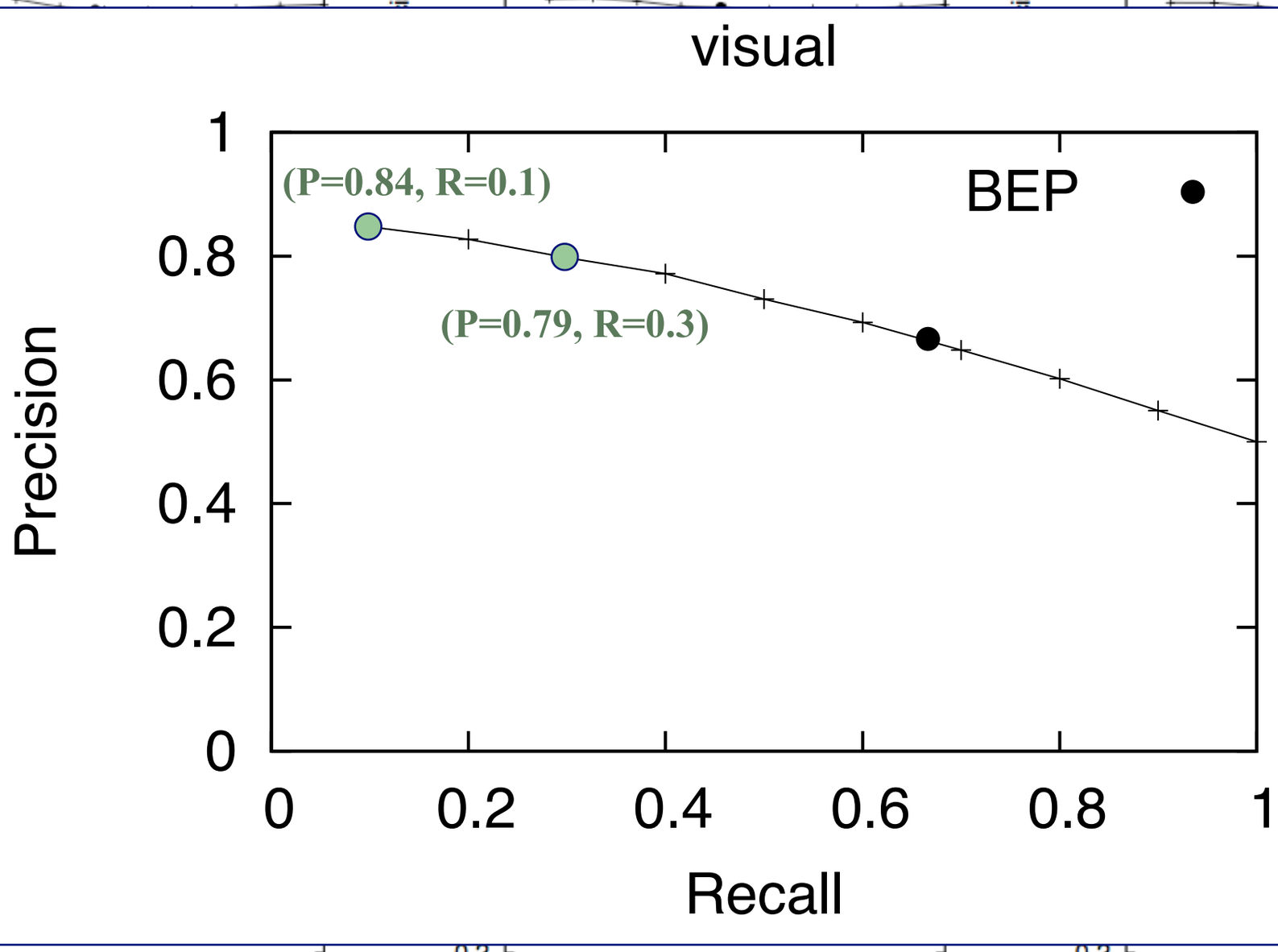


brightness

contrast

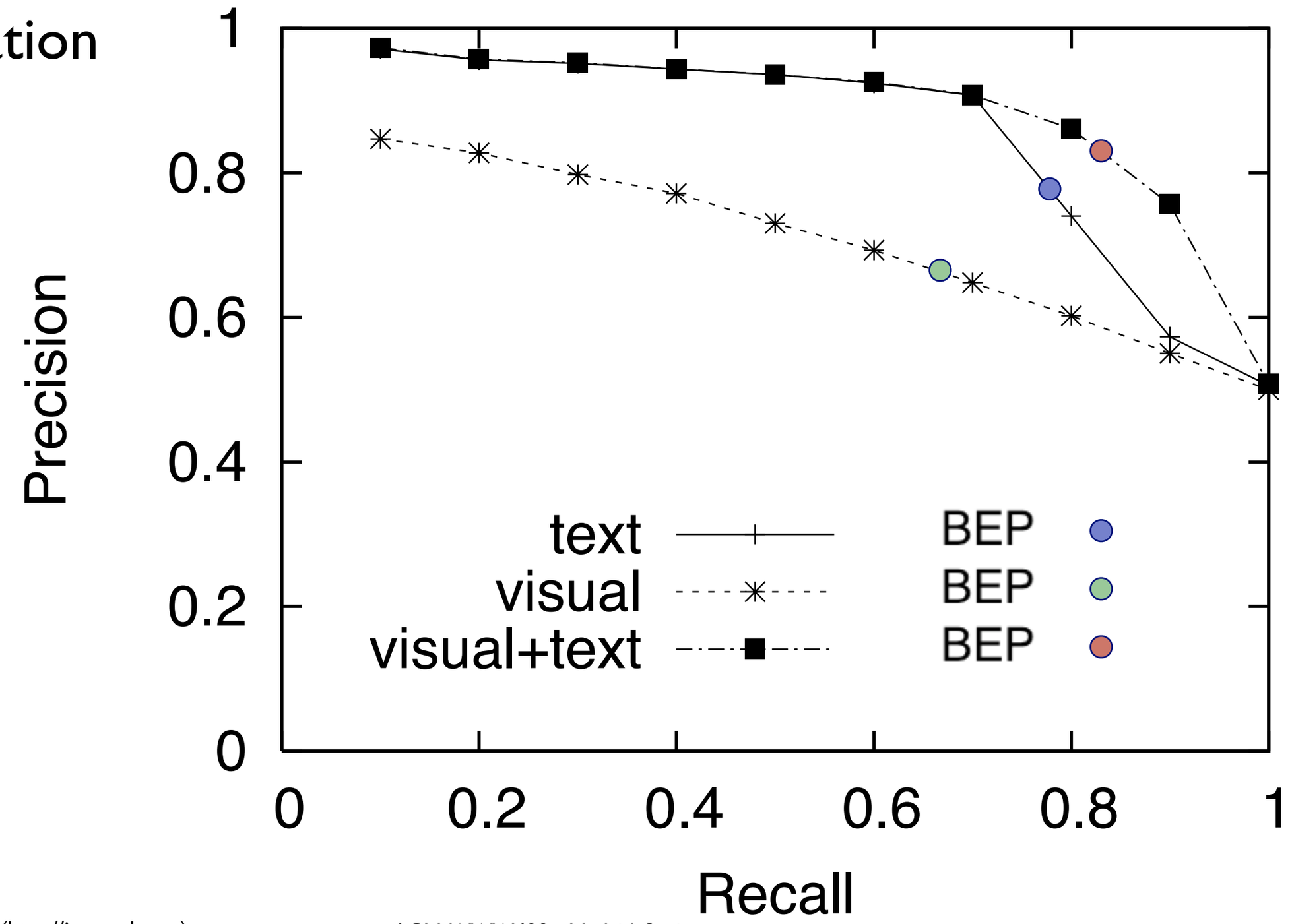
RGB contrast

visual



Experiments

Classification



Experiments

Ranking

- SVM Regression
- Training set:
  - 20,000 random photos with NumFav >= 2
  - 20,000 random photos with NumFav = 0
- Test set:
  - Remaining (disjoint) set
- Test set ranked and compared to decreasing NumFav sorted list
  - Kendall's Tau-b is used to compare ranked lists

Experiments

Ranking

Method	Kendall's Tau-b
brightness	0.0006
contrast	-0.0172
RGB contrast	0.0288
saturation	0.1064
saturation variation	0.0472
colorfulness	-0.0497
sharpness	0.0007
sharpness variation	-0.0914
naturalness	0.0143
text	0.3629
visual	0.2523
text+visual	0.4841

$$\tau_b \in [-1, 1]$$

Conclusions

Successful strategy to mine

- Community Feedback
- Metadata
- Visual Content features

Specifically

- Classification & Ranking models of image attractiveness using:
  - Textual features from meta data: Tags
  - Visual features from image content
- Ground-truth (Class labels and relevance values)
  - Community feedback: Number of favorite assignments

Conclusions

Results show

- Hybrid approach (text + visual) offers the best performance
  - High precision-recall - BEP = 0.8363, T=8000, N.Fav>=5
- Visual models provide applicable results
  - Lower BEP : 0.6664
  - Higher flexibility
    - Local domain
    - Recently updated pictures (no feedback still available)

Future Directions

- Development & Evaluation of enhanced photo search tools
- Extension to other kinds of media
  - Videos (YouTube)

Thank You