ETH zürich

Affective Image Classification using Features Inspired by Psychology and Art Theory

Jana Machajdik & Allan Hanbury, 2010

Alice Balest, Damian Gerber, Nicola Lo Russo

Overview

Main objective

- **Emotional Semantic Image Retrieval (ESIR):** Classify images into their perceived emotional content
- Find features in an image that correspond to a specific emotional content

Method

- Use concepts from psychology and art to define features that are potentially connected to emotional expression (e.g. color red to anger)
- Classification into **eight emotional categories:** Amusement, Awe, Contentment, Excitement, Anger, Disgust, Fear and Sad
- Determine the most relevant features

Background and importance

Motivation

- Huge and constantly growing image databases on the internet
- Need for **categorization based on emotional content** (e.g. journalism, AI generated images)
- Most applications lack the ability to incorporate the affective level
- Goal to **mimic human decisions**. Step towards training machines to human emotions.

Preliminary work

- Few publications in affective content analysis
- There is advanced work but it usually has
  - too common and generic features or
  - chosen emotional categories are too specific to the studies

Methodology

- **Preprocessing**
  - Image resizing and cropping
  - Color space conversion (RGB to cylindrical)
  - Waterfall segmentation algorithm
- **Feature extraction**
  - Define features
  - Compute features for each image
  - 114 total features

Preprocessing

- **Resize images**
  - To 200,000 pixels (maintaining aspect ratio)
  - Gain computational efficiency
- **Remove borders**
  - Hough transform + Canny edge detector
  - Crop single-color borders
- **Convert color space**
  - RGB to cylindrical coordinate color space
  - Intuitive definition of Hue (H), Saturation (S), and Brightness (Y) channels
- **Image segmentation**
  - Waterfall segmentation to get contiguous regions within images
  - Analyze spatial composition

Features

- **Feature categories**
  - Psychology-based features
    - Saturation and brightness related to emotional dimension
  - Art theory-based features
    - Relate color combinations to induced emotional effects
  - Standard features
    - Proven utility in image retrieval and classification tasks
- **Feature groups implemented**
  - Color
  - Texture
  - Composition
  - Content

Color features

- **Saturation** and **brightness**
  - pleasure, arousal and dominance
- **Hue** (tone)
- **Colorfulness**
  - with Earth Mover's Distance (EMD)
- **Color names** (number of pixels per color)
- **Itten contrasts** and color accordances
- **Wang's histogram**
  - emotional impact of color

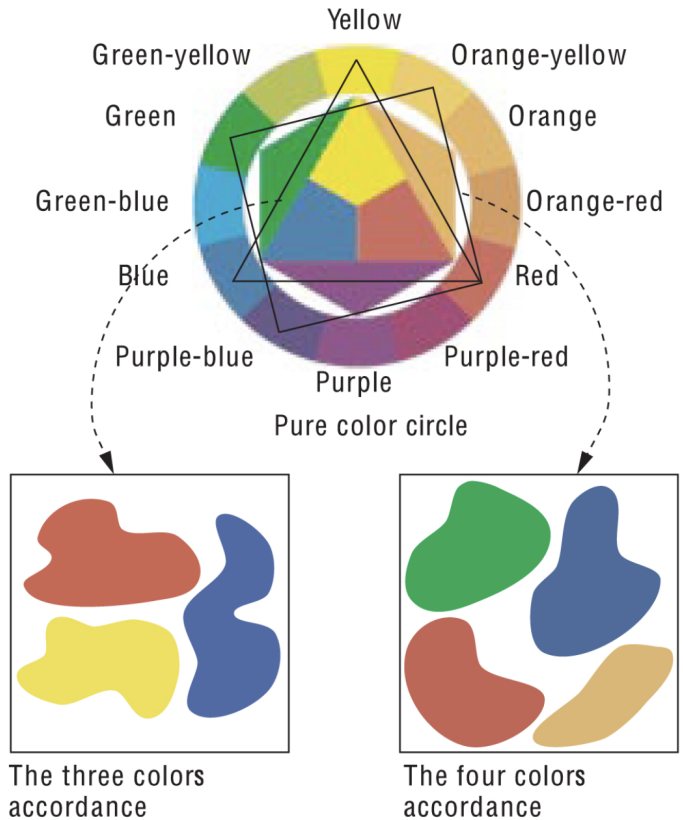


Figure: Color accordances.

Texture features

- **Wavelet-based** features
  - Measure spatial smoothness/graininess
  - Compute 9 wavelet features (H, S, Y) over 3 levels
- **Tamura** texture features
  - Coarseness, contrast, and directionality
- **Gray-Level Co-occurrence Matrix (GLCM)**
  - Compute contrast, correlation, energy, and homogeneity

Composition features

- **Level of Detail**
  - Impact of image detail
- **Low Depth of Field** (DOF)
  - Focus on the main subject
  - Blurring of the background
- **Dynamics**
  - Emotional effects of lines
  - Different lines convey different emotions
- **Rule of Thirds**
  - Divide the image into thirds, positioning the main object

Figure: Rule of Thirds applied to two images.

Content features

Semantic content has the greatest impact on emotional influence.

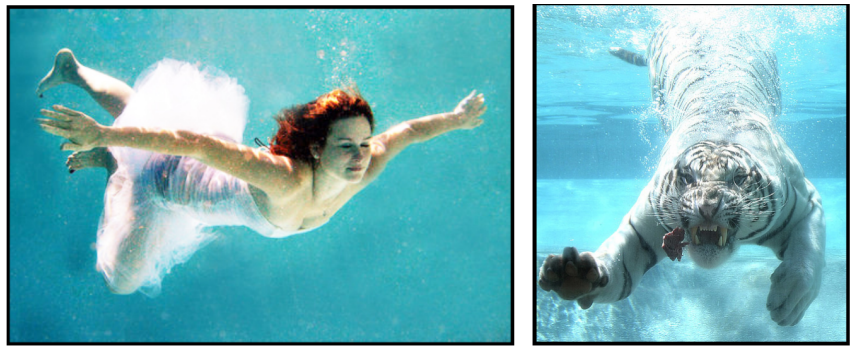


Figure: Two similar images with different contents.

- **Human faces**
  - Faces draw attention
  - Detect frontal faces using Viola and Jones' algorithm
  - Count faces and consider the size of the largest face
- **Skin detection**
  - Detect "artistic nudes"
  - Measure skin area and its proportion to detected face size

Experiment

Emotional Output Categories: Amusement, Awe, Contentment, Excitement, Anger, Disgust, Fear, Sadness (see psychological study on affective images).

Setup

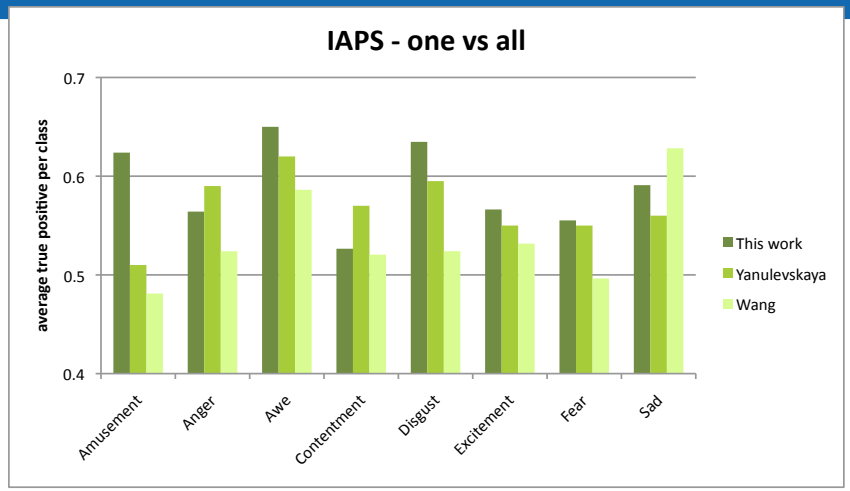
- Rule: “one category against all”
- Optimize for the true positive rate per class averaged over the positive and negative classes

Experiment - Testing and training data set

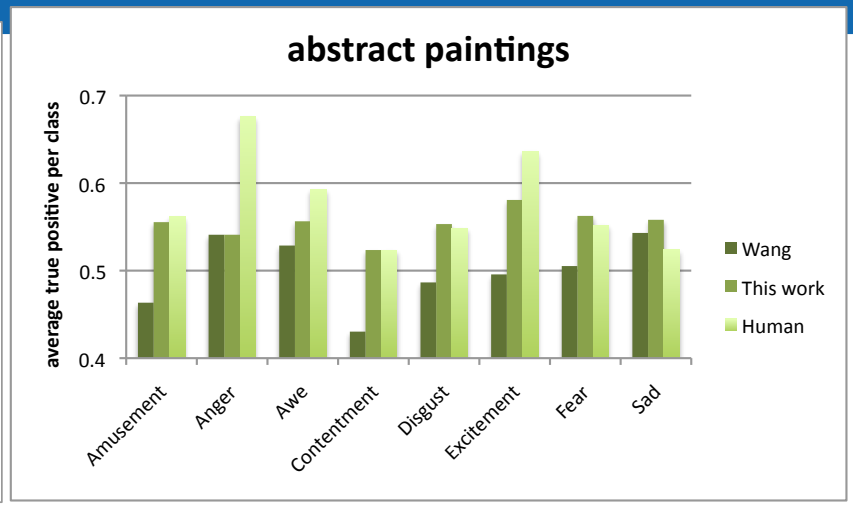
- A set of 394 images from **The International Affective Picture System (IAPS)**
  - Used in emotion research
  - Documentary natural color photos depicting objects
- A set of 807 **artistic photographs** downloaded from an art sharing site
  - Emotion determined by the artist
  - Conscious manipulation of the image composition, lighting, colors, etc
- A set of 228 peer rated **abstract paintings**
  - Combinations of color and texture
  - Each image rated about 14 times
- A combined set of all of the above

Classification performance results

- **IAPS**:
  - Best feature set is dependent on both the category and the data set
  - Occurrence and size of human faces is the strongest feature for the Amusement category
  - Categories are strongly content related
- **Artistic sets**
  - No strong connection between faces and categories
  - Colors are more important for photos (features from art theories (Itten colors, Wang Wei-ning histograms) are more common)
    - Itten features are selected for the Amusement and Excitement classes
    - Wang Wei-ning et al. features are selected most for Awe and Disgust classes (also selected for the Amusement and Excitement classes)



(a) Performance for IAPS.



(b) Performance of abstract paintings set.

Figure: Classification performance taking best features for each category against the best features from previous works

- This work's feature sets outperform results by Yanulevskaya and Wang Wei-ning features in case of IAPS set for five categories
- This work's feature set and humans obtained similar results in the abstract paintings data set (not in the anger class)

Classification performance for all image sets

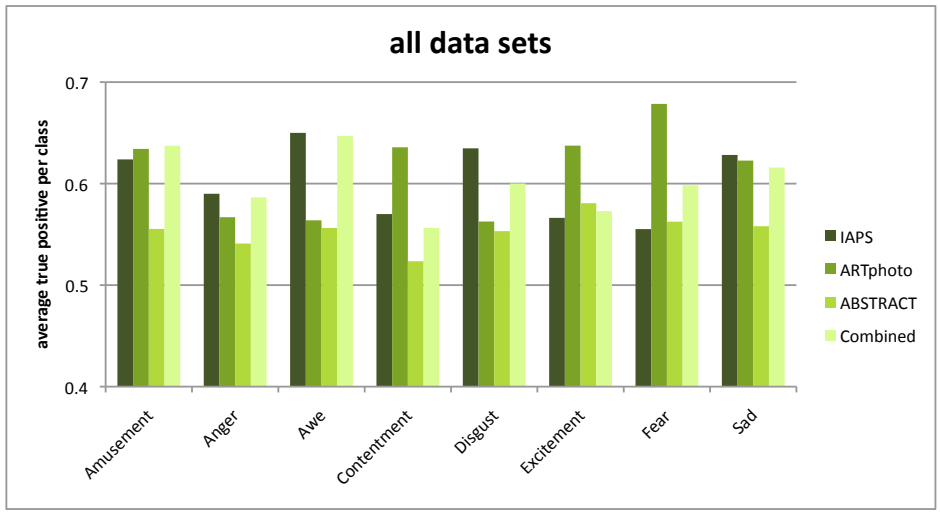


Figure: Classification performance for all image sets. Results are from the best feature selections

- No clear difference between IAPS and the artistic photo image sets
- Abstract paintings data set has the worst performance

Discussion and Conclusion

The algorithm achieved better results than the state of the art.

Critics: Cultural differences in how features and emotion are perceived could lead to false conclusions and discrimination.

It could be improved through

- New features (emotional expression, identify symbols (e.g. hearts, crosses))
- New Classification (distribution over the categories)
- More reliable ground truth (take larger samples of people)
- Different approach (drop consensus and learn individual preferences)
- Further research in psychology of affective images

Thank you!
Questions?

References

Jana Machajdik & Allan Hanbury, *Affective Image Classification using Features Inspired by Psychology and Art Theory*, 18th ACM international conference on Multimedia, 2010

Link to the pdf version

Category	Short Name	#	Short Description
color	*Saturation, Brightness*	2	mean saturation and brightness
	Pleasure, Arousal, Dominance	3	approx. emotional coordinates based on brightness and saturation
	Hue	4	vector based mean hue and angular dispersion, saturation weighted and without saturation
	Colorfulness	1	colorfulness measure based on EMD
	Color Names	11	amount of black, blue, brown, green, gray, orange, pink, purple, red, white, yellow
	Itten	20	average *contrast of brightness, contrast of saturation, contrast of hue, contrast of complements, contrast of warmth, harmony, hue count, hue spread, area of warm, area of cold,...*and the maximum of each
	Wang	19	features (histograms) by Wang Wei-ning et al. [31] (*factors 1 (10), factor 2 (7) and factor 3 (2)*)
	Area statistics	10	based on Wang features: *area of very dark, area of dark, area of middle, area of...light, very light, high saturation, middle saturation, low saturation, warm, cold*
texture	*Tamura*	3	features by Tamura et al [25]: *coarseness, contrast, directionality*
	Wavelet textures	12	wavelet textures for each channel (Hue, Saturation, Brightness) and each level (1-3), sum of all levels for each channel
	GLCM-features	12	features based on the GLCM: *contrast, correlation, energy, homogeneity* for Hue, Saturation and Brightness channel
composition	*Level of Detail*	1	number of segments after waterfall segmentation
	Low Depth of Field (DOF)	3	low depth of field indicator; ratio of wavelet coefficients of inner rectangle vs. whole image (for Hue, Saturation and Brightness channel)
	Dynamics	6	Line slopes: static, dynamic (absolute and relative), lengths of static lines, lengths of dynamic lines
	Rule of Thirds	3	mean saturation, brightness and hue of the inner rectangle
content	*Faces*	2	number of frontal faces, relative size of the biggest face
	Skin	2	number of skin pixels, relative amount of skin with respect to the size of faces

Figure: Features summary.

	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sad	sum
IAPS	37	8	54	63	74	55	42	61	394
Art photo	101	77	103	70	70	105	115	166	807
Abstract paintings	25	3	15	63	18	36	36	32	228
Combined	163	88	172	196	162	196	193	259	1429

Figure: Number of images per data set and emotion category.

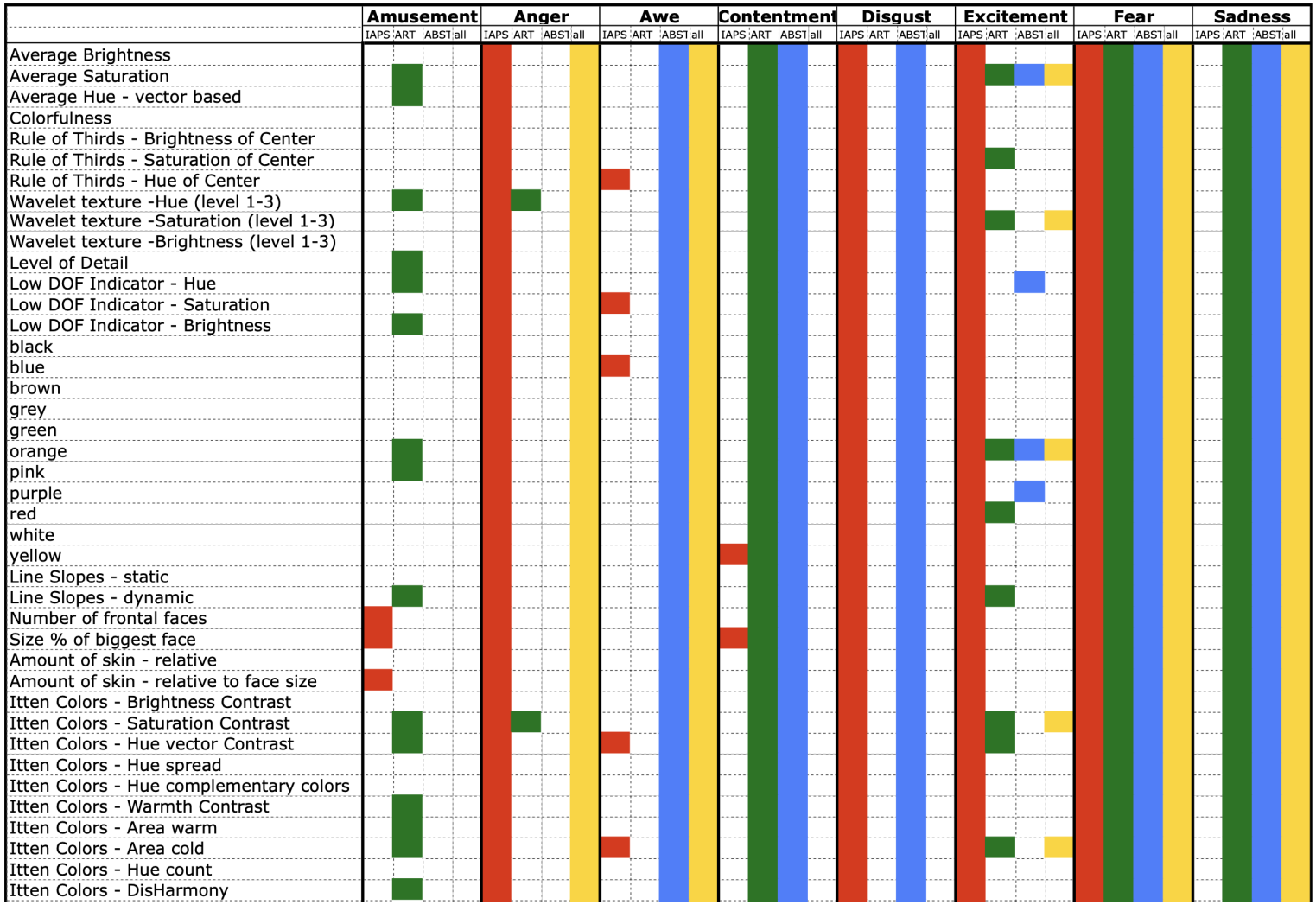


Figure: Detailed results (1/2).

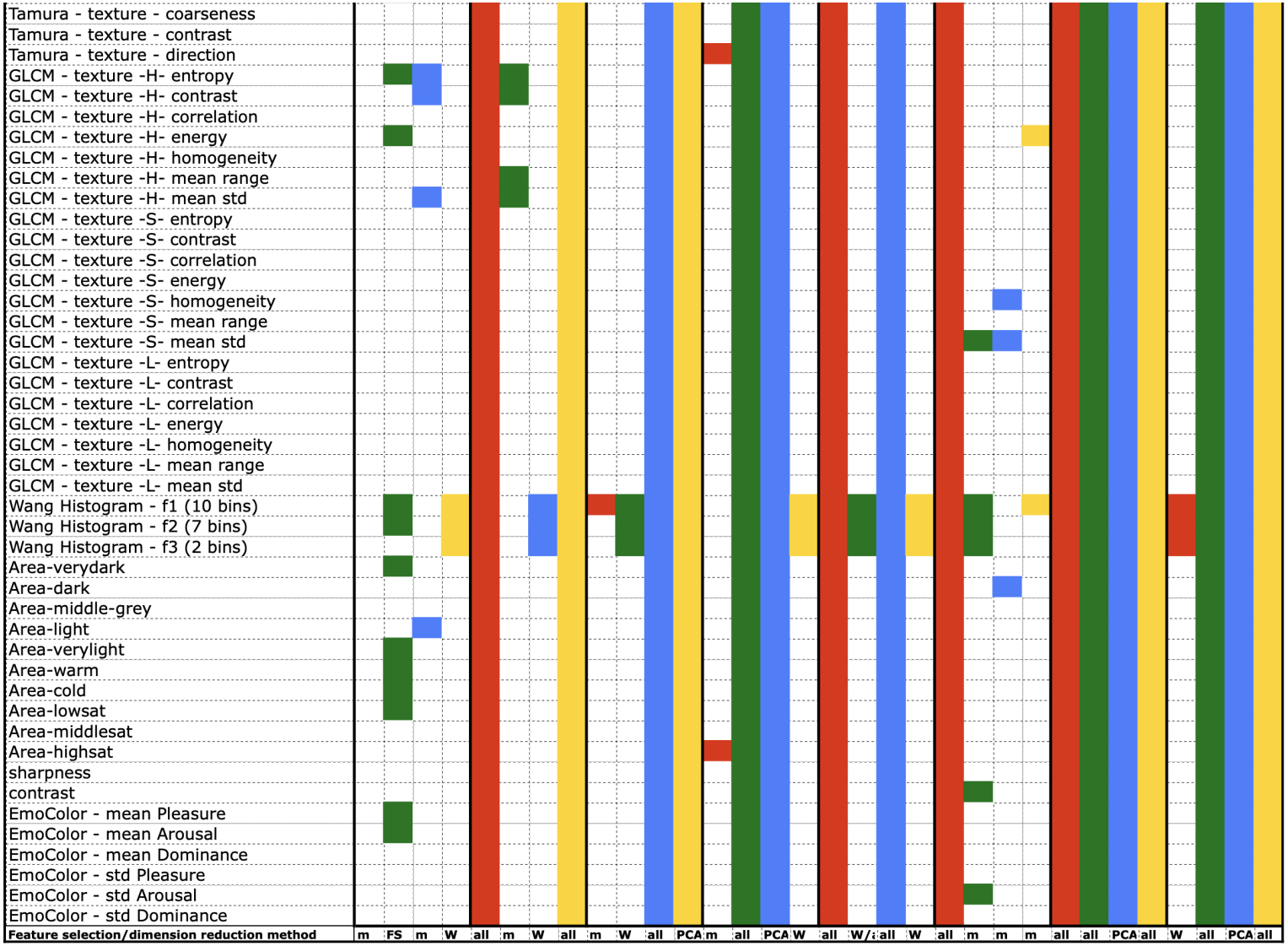


Figure: Detailed results (2/2).