

# The Flux: Creating a Large Annotated Image Database

Daniel Tamburrino, Patrick Schönmann, Patrick Vandewalle, and Sabine Süsstrunk

School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

## ABSTRACT

From image retrieval to image classification, all research shares one common requirement: a good image database to test or train the algorithms. In order to create a large database of images, we set up a project that allowed gathering a collection of more than 33000 photographs with keywords and tags from all over the world. This project was part of the “We Are All Photographers Now!” exhibition at the Musée de l’Elysée in Lausanne, Switzerland. The “Flux,” as it was called, gave all photographers, professional or amateur, the opportunity to have their images shown in the museum. Anyone could upload pictures on a website. We required that some simple tags were filled in. Keywords were optional. The information was collected in a MySQL database along with the original photos. The pictures were projected at the museum in five second intervals. A webcam *snapshot* was taken and sent back to the photographers via email to show how and when their image was displayed at the museum.

During the 14 weeks of the exhibition, we collected more than 33000 JPEG pictures with tags and keywords. These pictures come from 133 countries and were taken by 9042 different photographers. This database can be used for non-commercial research at EPFL. We present some preliminary analysis here.

**Keywords:** Image database, amateur photography, tags, image keywords, colorfulness, color naming

## 1. INTRODUCTION

Digital photography has revolutionized the way images are captured. Most people have a camera in their pockets, integrated into their cell phone or as a small device. As memory is extremely cheap, it is also hardly more expensive to take 100 than 10 pictures. “Think less, shoot more” could be today’s photography slogan.

The democratization of digital photography has changed not only amateur photography, but also professional photography. Amateurs sometimes enter areas that were until very recently limited to professionals. Everyone can be a witness of an exceptional event and take an exclusive picture with his mobile phone or digital camera. Even though the quality might be very poor, this picture could interest mainstream media. A well known illustration is the 9/11 terrorist attack or the London Underground bombing, from which amateur pictures and videos were diffused all around the world; on the Internet but also in newspapers or on television. Today, there exist web agencies such as *Scoop*<sup>1</sup> that are willing to buy amateur pictures to sell them to the media.

Analyzing this rapid evolution, the Musée de l’Elysée (Lausanne, Switzerland)<sup>2</sup> mounted an exhibition about the mutation of amateur photography in the digital age. This Swiss photography museum not only wanted to showcase amateur photography today, but also present its changing relationship with professional photography. This is how the “We are all photographers now!” exhibition<sup>3</sup> was born.

The huge amount of pictures taken every day does not only interest the artistic community. Regarding the increasing number of images available on the home PC or on the Internet, it becomes difficult to find the exact images we are looking for. A lot of research is done to study, analyze, classify, sort, or find images. Those research topics all share one common requirement: they need a good image database to test or train the algorithms. Such a database is often difficult to find: too small, too expensive, not corresponding to the researcher’s needs, etc. The best solution is thus to create one. However, using ones own pictures is not always recommended: the database may be influenced by the research hypothesis or may not contain a sufficiently large number of images.

---

Further author information:

E-mail: daniel.tamburrino@epfl.ch

Internet being the biggest source of images, it would be easy to download them and create a huge database. The problem with this approach is copyright, i.e. the lack of authorization to exploit those pictures. A community website like Flickr,<sup>4</sup> however, has millions of pictures, which are under common creative license and thus free to use, and sometimes even free to modify. Furthermore, Flickr offers the possibility for the photographer and for the visitor to add tags to photos. Downloading those images can be a good solution. The only drawback is that there is no control at all over the keywords and tags used for the pictures. They might not be appropriate or accurate. There is also the Google Image Labeler<sup>5</sup> project where people are asked to annotate images in order to improve Google Image search results. The difference again is that arbitrary viewers annotate the images, not the photographers themselves. Furthermore, those tagged images are not available to download and use.

## 2. THE FLUX PROJECT

The Flux project was born from the need for a good image database for research purposes. The project is a collaboration between the Musée de l'Elysée and the Image and Visual Representation Group at LCAV, EPFL (Switzerland). The goal was to collect a large number of photos with specific tags and some keywords. Every photographer, amateur or professional, could upload his pictures on a website. The incentive for the participants was that each uploaded picture was displayed in the Musée de l'Elysée during the exhibition "We Are All Photographers Now!". Moreover, a webcam capture of the displayed photo in its surrounding (i.e., an exhibition view, see Fig. 1) was taken and sent to the photographer via email. For many people, having a photo exposed in a museum, even if it is only for a few seconds, was sufficient motivation to upload a picture and fill in some tags and keywords.



Figure 1. Exhibition view taken in the museum by a webcam and sent back to the photographer via email.

To submit an image, we required that some simple fields (*title*, *subject*, *location*, *date*, and *category*) were filled in. Keywords were optional. It was also required to accept the terms and conditions, which authorize us to use the pictures for research purposes. The information along with the images allows us to study how people tag their own photos. The information can be very precise or completely subjective. For example, the tag *location* can contain all kinds of information, from *home* to a street address or even GPS coordinates.

Several steps were needed to set up such a project. First, a website was created where people can upload their photos. The pictures and data had to be stored in a database. As the image size could be as much as 6Mb, the files were subsampled (or upsampled in the case of some camera phone images) to adapt them to the projection dimensions, in order to reduce bandwidth usage. The pictures were then manually validated before they were made available for display in the museum. The photos were pseudo-randomly selected, such that each picture was displayed at least once. Next, we had to synchronize the webcam capture with the photo projection. Finally, the exhibition view had to be sent back to the author via email.<sup>6</sup>

## 2.1 Backend

The backend was composed of two systems. The first one, at EPFL, hosted the allphotographersnow.ch website along with a MySQL database. The second system was located at the museum. It handled the slideshow and controlled the Ethernet webcam that captured the exhibition views. This was all done using a single webpage, running under Firefox. An AJAX request periodically downloaded a batch of pictures from the main server and displayed them for 5 seconds each with the name and country of the photographer in overlay. Each time an image was displayed, a request was sent to the webcam to capture the exhibition view. The snapshot was associated with the photo and photographer's email using a timestamp. Once a week, the exhibition views were sent out to the photographers via email. This was done when the museum was closed and the slideshow was therefore not running, in order to reduce the server load.

Is five seconds per image enough? If the images would have been displayed longer, the visitors could easily get bored. If various pictures are refreshed at a sufficient rate, the visitors are kept interested and want to watch more. At this rate, more than 5000 pictures were displayed every day during opening hours. In our experience, the audience was thrilled. Some people would stay hours to look at the continuous flux of pictures, to discover the next one, maybe secretly hoping to see their images on the wall.

The screenshot shows the 'All Photographers Now! - Administration zone' in Mozilla Firefox. The browser address bar shows 'http://www.allphotographersnow.ch/admin/'. The page features a large red banner at the top left that says 'WE ARE ALL PHOTOGRAPHERS NOW!'. Below the banner, there are several sections:

- General statistics:** Photos: 33'881, Photographers: 9'042, Countries: 133.
- Time statistics:** Days since the upload site is functional: 123, Days since the beginning of the exposition: 101, Days left for the exposition: 0.
- Average statistics:** Photos per photographer: 3.75, Photos per day (yesterday): 0, Photos per day (last 7 days): 55, Photos per day (from the beginning): 275.
- Min/Max statistics:** Maximum uploads in one day (2007-02-08): 1'466, Minimum uploads in one day (2007-05-22): 5, Maximum uploads for one photographer: 153.
- Predictions:** Estimated final number of photos: ~33'900.
- Flux display statistics:** Average: 12, Standard deviation: 11, Minimum: 0, Maximum: 65.

The main area of the page displays a list of photos for moderation. Each photo entry includes a thumbnail, a 'details' link, and a table of 'UPLOAD INFORMATION'. The table columns are 'Submitted', 'By', and 'As'. The photos listed are:

- PHOTO 19'618:** Submitted 78 days ago, By David Martel, Guernsey@gmail.com, As DSCF0023\_sharpen.jpg. Title: Creature of the Night.
- PHOTO 19'619:** Submitted 78 days ago, By GERARD DECOQ, France, As La\_madone\_.jpg.
- PHOTO 19'620:** Submitted 78 days ago, By Carole Fischer, Switzerland, As Panorama1.jpg.
- PHOTO 19'621:** Submitted 78 days ago, By Beatrice Jann, Switzerland, As DSC00736\_3\_1.jpg.
- PHOTO 19'622:** Submitted 78 days ago, By Belén Carrasco-Muñoz Ortiz, Spain, As Compa\_eros.jpg.
- PHOTO 19'623:** Submitted 78 days ago, By Belén Carrasco-Muñoz Ortiz, Spain, As California.jpg.

Figure 2. Moderation webpage used to authorize or ban pictures.

## 2.2 Image Authorization

The exhibition was intended for a very large audience, including children. We therefore decided to ban pornography, strong violence, or any inappropriate content. Image authorization was done manually. The validation process used AJAX requests to dynamically update the webpage (Fig. 2) as the reviewer authorized or banned pictures. Duplicates were marked using a function that simply looked at the name and size of the uploaded files for a given photographer's email address.

Out of the more than 34000 uploaded pictures, only 482 were banned, i.e. 1.5%. This was less than expected. Furthermore, 90% of the banned images were duplicates: people sent the same picture twice or more. Less than 50 pictures were pornographic or with an inappropriate content. We also received some images that were not photos, but screenshots of Second Life's characters. It is interesting to observe that some people start to associate these screen captures with photos, and thus the virtual world with the real world.

## 3. EXHIBITION STATISTICS

During the 14 weeks of the exhibition, from February to May 2007, we collected more than 33000 unique JPEG images with tags and keywords. All information provided by the photographer was stored in a MySQL database along with the original picture. The latter might also contain EXIF data, if the editing process did not remove it. These pictures come from 133 countries and were taken by 9042 different photographers. The main contributors are from Switzerland (30%), USA (14%), France, UK, Italy, and Germany (Fig. 3). Each photographer sent an average of 3.75 pictures. Although it was not mandatory, 69% of the pictures came with keywords, with an average of 3.3 keywords per picture.

The daily number of uploaded pictures varies considerably (Fig. 4). Favorable newspaper and TV coverage rapidly spread the word about the website allphotographersnow.ch, not only in Switzerland, but all around the world. The high peaks correspond to newspaper or TV coverage. The daily rate of uploaded pictures progressively decreased over time. It is nevertheless interesting to note that the rate increased slightly every Monday, when emails with the exhibition view were sent.

Another interesting finding is the date when the pictures were taken (Fig. 5). The uploaded pictures are quite recent. 74% of the pictures were taken within the last 1.5 years, and 87% within the last 2.5 years. Does this mean that people prefer their latest pictures, or that they simply did not have a digital camera before? It would certainly be interesting to study this further, but it is beyond the scope of our current research.

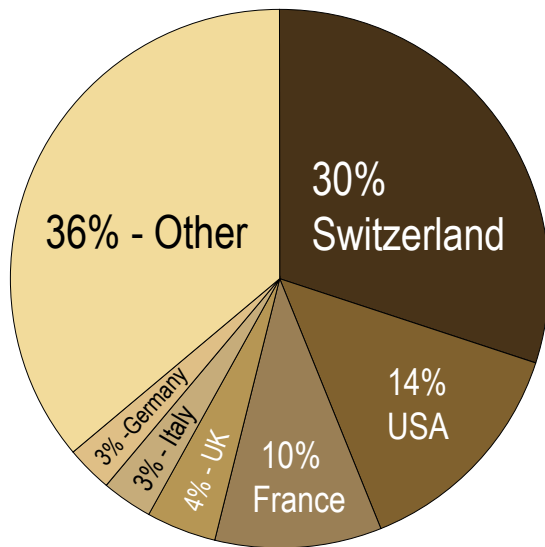


Figure 3. Countries of origin of the uploaded pictures.

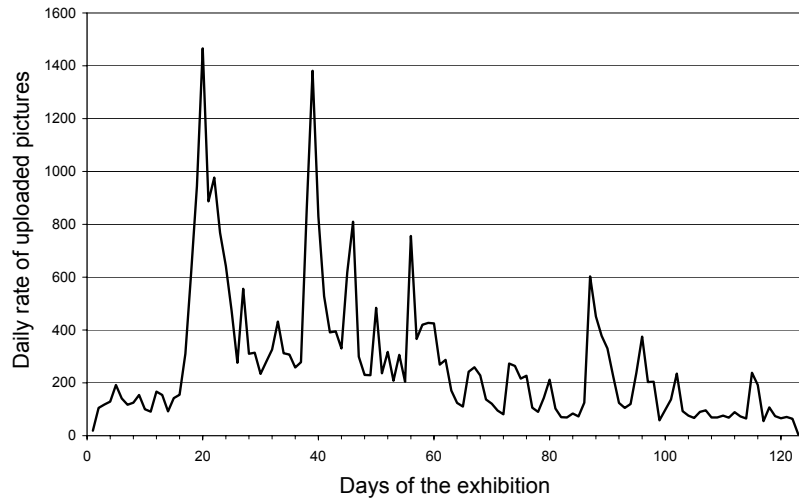


Figure 4. Rate of photos uploaded per day for the duration of the exhibition.

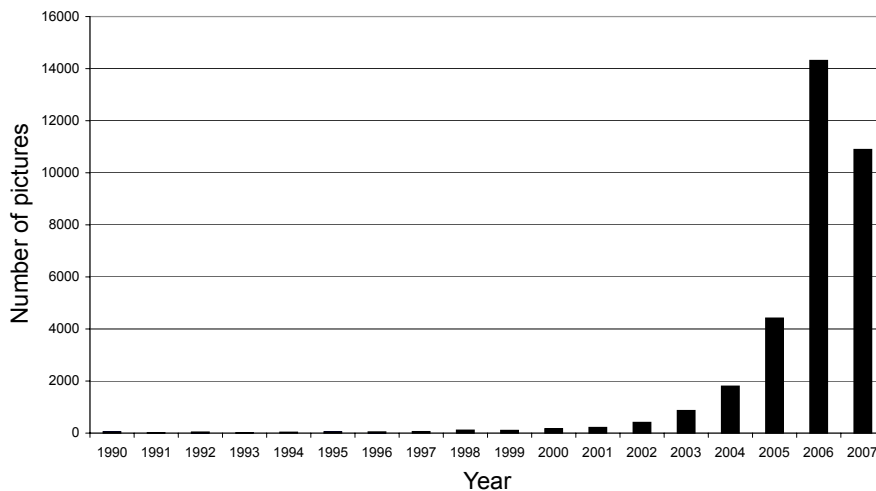


Figure 5. Year in which the uploaded pictures were taken.

#### 4. KEYWORDS

All kinds of information can be extracted from the huge amount of data that come along with the pictures. We are mostly interested in people's relationship with their photos. The scope of this section is to give an intuitive view of what kind of keywords people associate with their pictures. This is only a preliminary report, we are still continuing the evaluation.

We analyzed the most often occurring keywords and categorized them into logical categories, along with similar keywords. Note that we are looking only at keywords, not at the tags. Each keyword is shown with the number of its occurrence in the database. The different spellings (e.g. color or colour), plural form, and French translation are taken into account even though they are not explicitly mentioned in the tables.

The different groups of keywords are listed in Tables 1 and 2.

Analyzing the different keywords and their occurrences, we can draw the following conclusions. Photographers tend to use more often general keywords, rather than specific ones. They give general information about the content of the picture without going into details, as illustrated in the *image element* class (Tab.1):

Image type		Image element		Time		Color	
Keyword	Counts	Keyword	Counts	Keyword	Counts	Keyword	Counts
portrait	1088	water	997	summer	716	color	868
art	883	sky	871	night	596	blue	649
landscape	688	sun	645	winter	465	red	495
architecture	478	beach	496	spring	242	black and white	473
		snow	462	fall	192	green	371
		sunset	410	morning	155	yellow	237
		building	303	day	137	orange	190
		house	287	evening	104	pink	143
		ocean	196	afternoon	24	brown	72
		sand	152			purple	66
		fire	146			violet	17
		grass	137			magenta	4
		flame	19			cyan	1
		cloud	6				

Table 1. Number of images that have the specified keyword from the specified class.

People (objective)		People (relative)		Animals		Event	
Keyword	Counts	Keyword	Counts	Keyword	Counts	Keyword	Counts
child	656	friend	135	animal	349	holidays	151
woman	556	son	63	cat	231	war	100
girl	502	daughter	54	dog	212	christmas	79
man	386	sister	36	bird	191	wedding	52
boy	213	brother	33	puppy	18	funerals	26
baby	102	wife	13	kitten	14		
female	72	husband	11				
male	33	uncle	4				
		aunt	1				

Table 2. Number of images that have the specified keyword from the specified class.

fire(146)  $\gg$  flame(19), beach(496)  $\gg$  sand(152), and sky(871)  $\gg$  cloud(6). Obviously, there are flames in fire, sand on the beach, and sometimes clouds in the sky. Another example is found in the *time* class (Tab.1): seasons are much more used than the time of the day (morning, afternoon, evening). Note also that more pictures seem to have been taken in summer (716) and winter (465) than in spring (242) and fall (192). Of course this is subject to interpretation, as photographers might advantage one season when tagging their images, as it occurs with day (137) and night (596): there are more day than night pictures, but there are more *night* than *day* keywords, as people consider a day picture more usual than a night shot.

One of the most common photo types is portrait. People often appear in images. Again, photographers use more often general keywords than specific ones. Table 2 shows that objective words describing people (man, boy, etc.) are more used than relative terms (son, brother, etc.).

Color keywords (Tab.1) are quite often present. Interestingly, their occurrence closely follows the color term order established by Berlin and Key,<sup>7</sup> with the exception of blue, which is more frequently used than the other color names. One can speculate that blue scene elements, such as water or sky, are more often present in images as opposed to our visual consciousness. Some color names familiar to the technical imaging crowd, such as cyan and magenta, are not part of photographers' color vocabulary and occur very rarely.

The *color* (868) and *black and white* (473) keywords are also often used. Although there is twice as much occurrences of the *color* keyword compared to *black and white*, the image database contains 91% of color images. The *black and white* keyword is thus proportionally much more often indicated than *color*. As for the *night* and *day* keywords, they illustrate that photographers use keywords to point out a non usual element in the photo, in this case a black and white picture in opposition to a color picture. The inverse is not true, *color* is not used

Attribute	$M$
not colorful	0
slightly colorful	15
moderately colorful	33
averagely colorful	45
quite colorful	59
highly colorful	82
extremely colorful	109

Table 3. Correspondence between the colorfulness metric  $M$  and the colorfulness attributes.

in opposition to *black and white*, but to mark the higher than average colorfulness of the image, as explained in the following section.

## 5. COLORFULNESS OF “COLOR” IMAGES

In this section, we show an example of the kinds of analysis our database allows. Among the different groups of keywords presented in section 4, the *color* category is best suited for a simple computational study, as it is a measurable parameter in the image. In this section, the colorfulness of images that contain the *color* keyword is compared to the colorfulness of images that contain at least one keyword, but not *color*. All computation is done on thumbnails images of size  $96 \times 72 = 6912$  pixels, which is sufficient for our needs.

In 2003, Hasler and Ssstrunk defined a metric for the colorfulness of natural images<sup>8</sup> by fitting a model to psychophysical experiments where subjects were asked how colorful an image was. The metric uses a very simple opponent color space and has a correlation of 95.3% with experimental data. Assuming that the image is coded in the sRGB color space, the opponent color space is defined as

$$\begin{aligned} rg &= R - G \\ yb &= \frac{1}{2}(R + G) - B, \end{aligned} \quad (1)$$

and the metric  $M$  is defined as

$$\begin{aligned} M &= \sigma_{rgyb} + 0.3\mu_{rgyb}, \\ \sigma_{rgyb} &:= \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2}, \\ \mu_{rgyb} &:= \sqrt{\mu_{rg}^2 + \mu_{yb}^2}, \end{aligned} \quad (2)$$

where  $\sigma$ . and  $\mu$ . are the standard deviation and the mean value of the pixel cloud along direction  $(\cdot)$ , respectively. The correspondence between the colorfulness metric and the colorfulness attributes is shown in Table 3.

Images with the *color* keyword (868 images) have a mean colorfulness value of  $\mu = 57.67$ , which correspond to the *quite colorful* attribute (Fig.6(a)). The median is 52 and the standard deviation  $\sigma = 32.53$ . Pictures without the *color* keyword, but with at least one other keyword (22528 images) have a mean colorfulness value of  $\mu = 35.74$ , which correspond to the *moderately colorful* attribute (Fig.6(b)). The median is 31.8 and the standard deviation  $\sigma = 26.20$ .

From the above results, photographers seem to associate the *color* keyword to more colorful images than those with average colorfulness. However, the standard deviation is quite large. This means that individual color objects in an otherwise more monochromatic scene are also considered as colorful. People do not use *color* as opposition to *black and white*, but to qualify the colorful colors in the picture. The inverse is not true: the *black and white* keyword is mostly used for black and white pictures, as shown in Fig.7 where the mean colorfulness value is  $\mu = 2.78$ , i.e. *not colorful*. The median is 0, and the standard deviation  $\sigma = 8.80$ .

Knowing that images with the *color* keyword are more colorful than images without this keyword, a further step is to determine which colors are present and in which proportions in those images. We therefore studied the hues present in these images, but also in images which have color names as keywords, such as red, green, blue, etc. The images were converted to the HSV color space. The hue plane was separated into six arbitrary regions, that are red, orange, yellow, green, blue, and purple (violet) (Tab.4). More specific or precise names

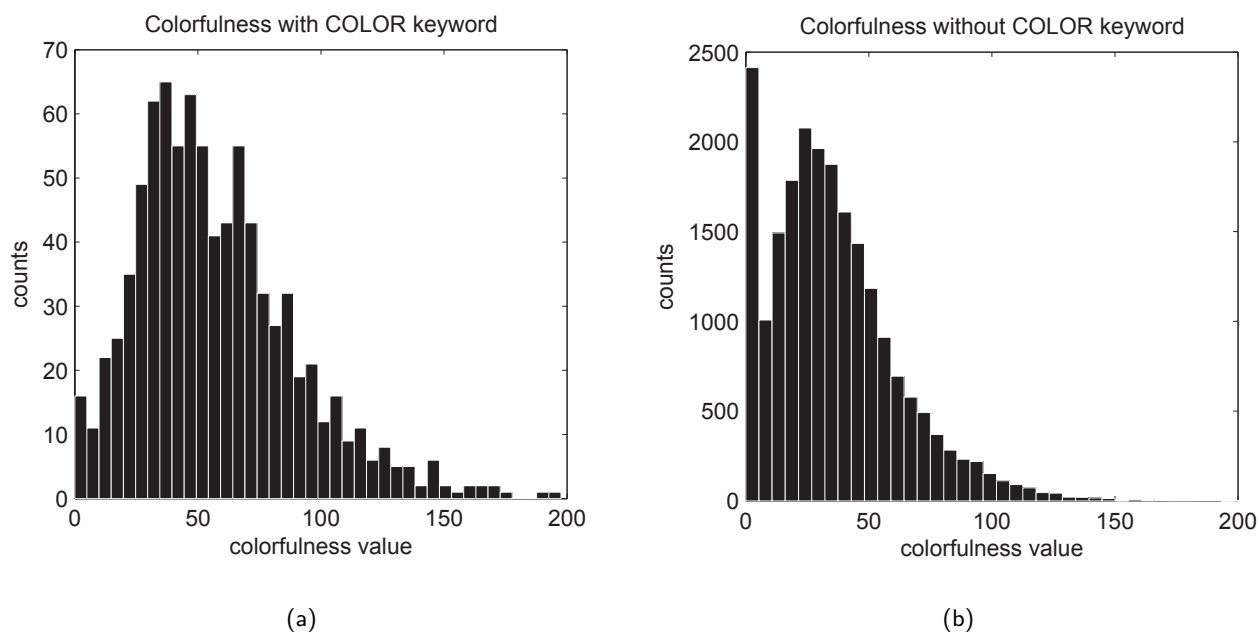


Figure 6. (a) Colorfulness of images with the *color* keyword. Mean  $\mu = 57.67$ , and standard deviation  $\sigma = 32.53$ . (b) Colorfulness of images without the *color* keyword. Mean  $\mu = 35.74$ , and standard deviation  $\sigma = 26.2$ .

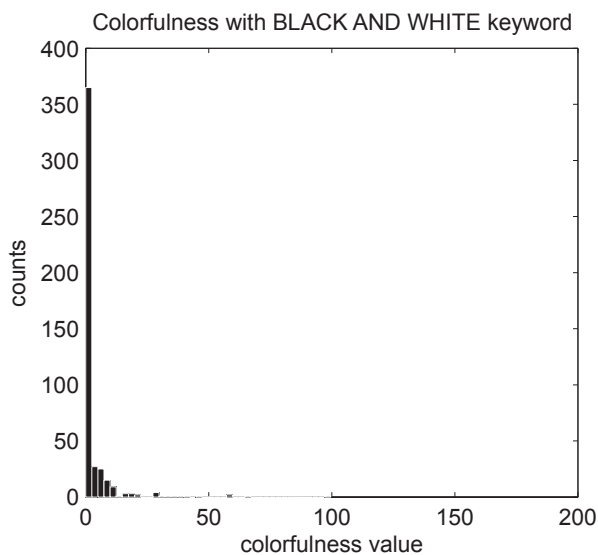


Figure 7. Colorfulness of images with the *black and white* keyword. Mean  $\mu = 2.78$ , and standard deviation  $\sigma = 8.80$ .

were not considered, as people do rarely use such color names, as seen in section 4 with the *cyan* and *magenta* keyword (Tab.1). The proposed algorithm counts the number of pixels in each hue region, discarding pixels for which saturation  $S$  and value  $V$  were below fixed thresholds. By setting the saturation threshold to  $S = 0.25$  and the value threshold to  $V = 0.3$ , too dark and not enough saturated pixels were not taken into account.

It appears that the distribution of hues in images with and without the *color* keyword are quite similar (Fig.8).



Hue name	Angle [deg]
Red	298 to 24
Orange	24 to 49
Yellow	49 to 66
Green	66 to 173
Blue	173 to 268
Purple	268 to 298

Table 4. Definition of the hue angles  $H$  of HSV and their corresponding color names.

Images not labeled as *color* are slightly more orange than red. Blue is strongly present in all images, and might correspond to the sky or water. In images with a specific color name as keyword (Fig.9), the dominant hue corresponds to the color name keyword for red, green, and blue. For orange, yellow, and purple, the proportion of those three hues is respectively bigger than in other categories, even though they are not the dominant color in the image. The keyword might have been attributed based on a small colorful object in the picture, not on the overall color of the image. For the *color* keyword more specifically, the result also depends on our definition of the red/orange region, which is subject to discussion. To do this rigorously, psychophysical experiments should be considered.

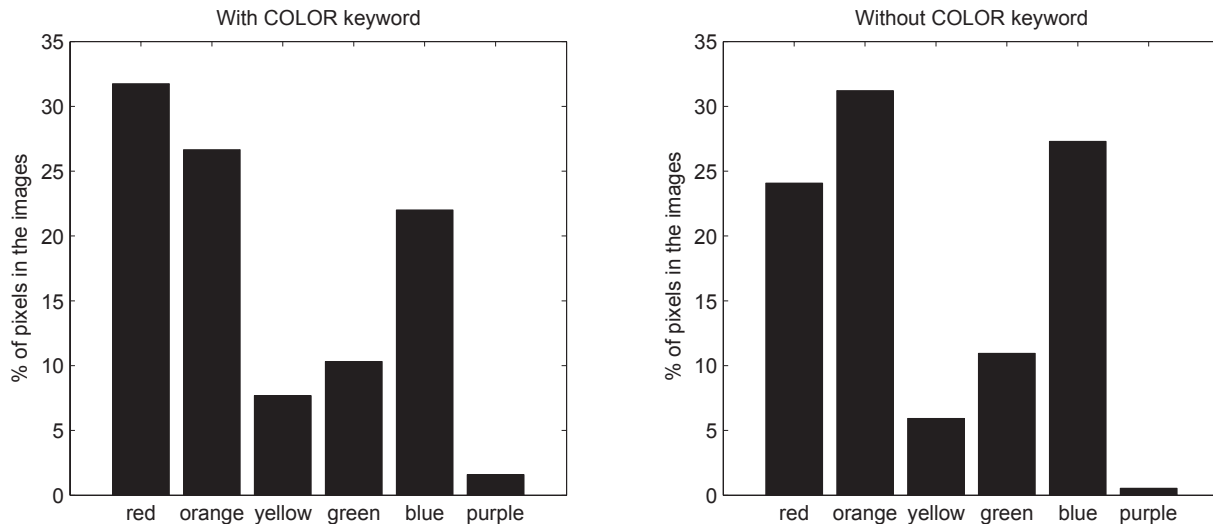


Figure 8. Distribution of hues in images with and without the *color* keyword.

## 6. CONCLUSION

The Flux project, and the entire “We Are All Photographers Now!” exhibition, was a success: one of the most - if not the most - visited exhibition at the Musée de l’Elysée. People liked it very much, and talked about it in their blogs, sharing the exhibition views of their photos. A simple query on Flickr with the keywords *allphotographersnow* and *tousphotographes* (French name of the exhibition) gives more than 1300 results.

From a research point of view, we have already started several projects. We are interested in knowing what kind of keywords people use. How can we associate those keywords to optimize image classification or image search? The state of the art algorithms can now extract objective features from an image, i.e. analyze it and extract keywords like *blue sky*, *portrait*, or *sunset*. Are these keywords the same as those used by humans to annotate their own images? Why are some images annotated as sunset not classified as sunset by our automatic classification algorithm? Having this database at our disposal will help us answer these questions.

## ACKNOWLEDGMENTS

The work presented in this paper was supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

## REFERENCES

1. <http://www.scoopt.com>.
2. <http://www.elysee.ch>.
3. <http://www.allphotographersnow.ch>.
4. <http://www.flickr.com>.
5. <http://images.google.com/imagelabeler/>.
6. D. Tamburrino, P. Schönmann, P. Vandewalle, and S. Süsstrunk, "Tous photographes! côté labo," *Flash Informatique Spécial été, EPFL*, 2007.
7. B. Berlin and P. Kay, *Basic Color Terms: Their Universality and Evolution*, Cambridge University Press, 1999.
8. D. Hasler and S. Süsstrunk, "Measuring Colourfulness in Natural Images," in *IS&T/SPIE Electronic Imaging 2003: Human Vision and Electronic Imaging VIII*, **5007**, pp. 87–95, 2003.

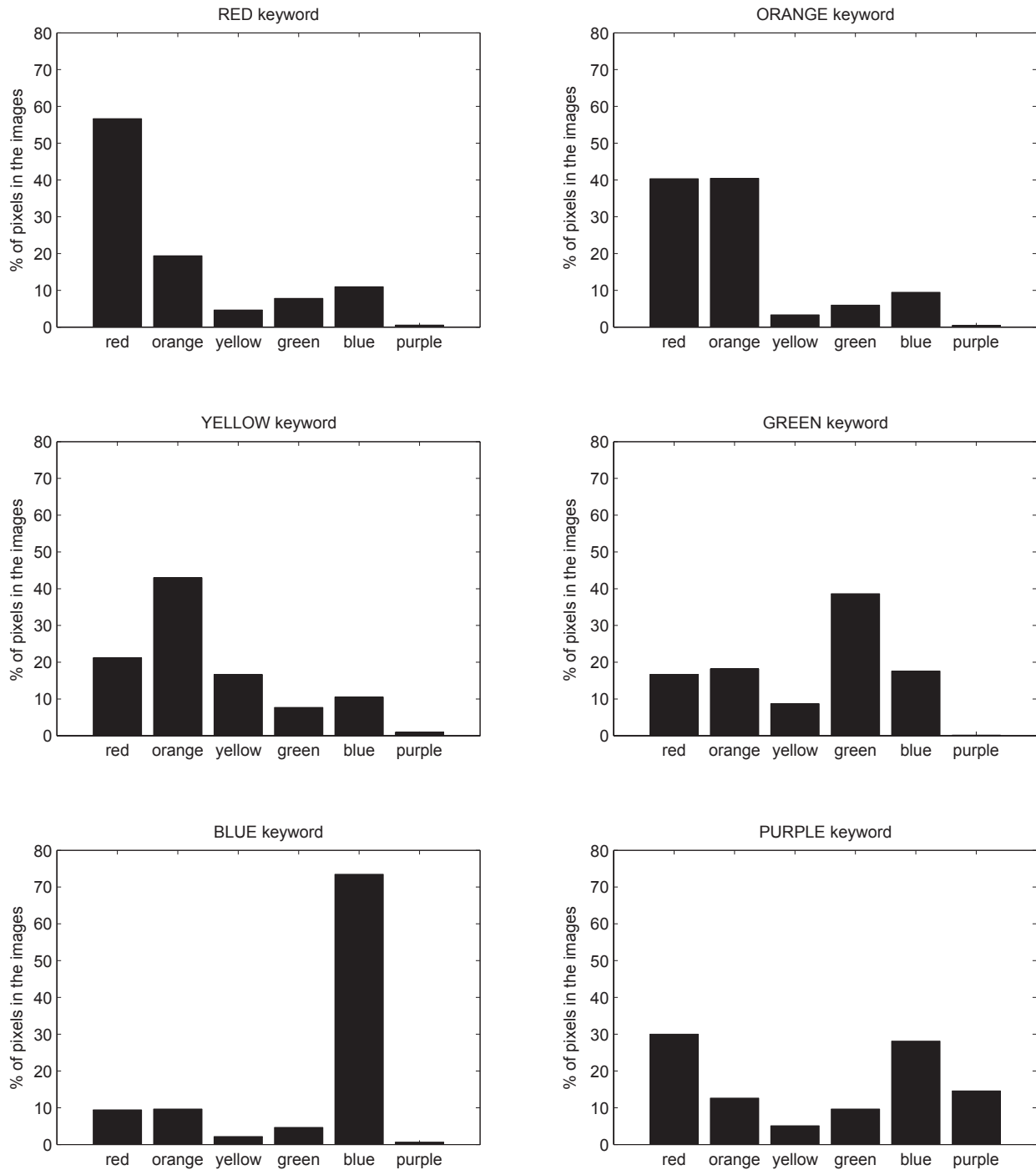


Figure 9. Distribution of hues in images with the keywords *red*, *orange*, *yellow*, *green*, *blue*, and *purple*, respectively.