How Shell and Horn make a Unicorn: Experimenting with Visual Blending in Emoji

João M. Cunha, Pedro Martins, Penousal Machado
CISUC, Department of Informatics Engineering
University of Coimbra
{jmacunha,pjmm,machado}@dei.uc.pt

Abstract

Emoji are becoming increasingly popular, both among users and brands. Their impact is such that some authors even mention a possible language shift towards visuality. We present a Visual Blending-based system for emoji generation, which is capable of representing concepts introduced by the user. Our approach combines data from ConceptNet, EmojiNet and Twitter's Twemoji datasets to explore Visual Blending in emoji generation. In order to assess the quality of the system, a user study was conducted. The experimental results show that the system is able to produce new emoji that represent the concepts introduced. According to the participants, the blends are not only visually appealing but also unexpected.

Introduction

The word *emoji* has a Japanese origin, in which the *e* means "picture", *mo* means "writing" and *ji* means "character"[1] – leading to the often attributed meaning "picture-word". Emoji seems to have become an important part of our way of writing. Their increasing usage is well documented by the importance given to them by language related resources – Oxford Dictionaries named the emoji "Face With Tears of Joy" the Word of The Year of 2015[2] – and by statistical data – Facebook reported in 2017 that 60 million emoji are used every day on Facebook and 5 billion on Messenger[3].

Some authors even discuss a shift towards a more visual language (Lebduska 2014; Danesi 2017). This shift would in fact bring us close to old ways of writing, such as hieroglyphs. Using images as complementary signs in written communication enriches it (Niediek 2016) by allowing the transmission of non-verbal cues (e.g. face expressions, tones and gestures) (Hu et al. 2017), which are lacking in written communication and Computer-Mediated Communication (CMC). This integration in written language is easy to observe when we consider the increasing number of emoji-related tools and features. Some examples are Search-by-
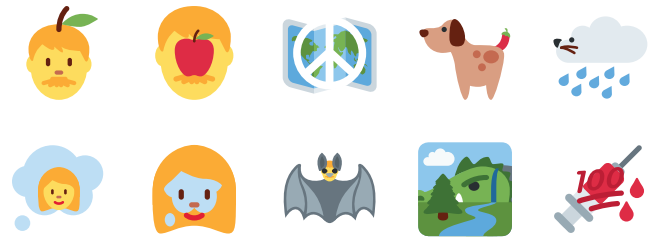

Figure 1: Examples of visual blends. From left to right and top to bottom: *man apple*, *apple man*, *world peace*, *hot dog*, *rain cat*, *woman wonder*, *wonder woman*, *man bat*, *dinosaur park*, and *true blood*

emoji supported by Bing[4] and Google[5], and the Emoji Replacement and Prediction features available in iOS 10[6]. We believe that other possible applications exist, specially in the domain of image generation (see some examples in Fig. 1).

Before emoji, sequences of ASCII characters were often used to express emotions CMC – emoticons (see Fig. 2). Despite the high adoption of emoji, some emoticons still continue to be used as an alternative due to their potential for customisation (Guibon, Ochs, and Bellot 2016). Whereas emoticons are composed of individual and replaceable parts, emoji are inserted as a whole in the text (Dürscheid and Siever 2017). In 2015, "skin tone" modifiers were added to Unicode core specifications and in 2016 the Unicode Consortium decided to implement the ZWJ (Zero-Width-Joiner) mechanism – an invisible character to denote the combination between two characters (Abbing, Pierrot, and Snelting 2017). This meant that new emoji could be created through the combination of existing ones, without the need to go through the Unicode Consortium.

Having the modifiers, the ZWJ mechanism and emoticons' combinational character as inspiration, it is our belief that Visual Blending can be explored to further extend emoji system. Visual Blending, which draws inspiration from Conceptual Blending (CB) theory (Fauconnier and Turner

---

[1] unicode.org/reports/tr51/proposed.html, retr. 2018

[2] en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2015,retr. 2018.

[3] blog.emojipedia.org/5-billion-emojis-sent-daily-on-messenger/, retr. 2018.

[4] blogs.bing.com/search/2014/10/27/do-you-speak-emoji-bing-does, retr. 2018.

[5] forbes.com/sites/jaysondemers/2017/06/01/could-emoji-searches-and-emoji-seo-become-a-trend/, retr. 2018.

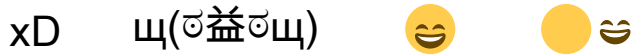[6] macrumors.com/how-to/ios-10-messages-emoji/, retr. 2018.

xD   щ(ಠ益ಠщ)   😄   🟡😄

Figure 2: "Laughing" western emoticon, "Why" eastern emoticon, "Grinning Face With Smiling Eyes" emoji, and deconstruction of an emoji

2002), is a Computational Creativity (CC) technique which consists in merging two or more visual representations (e.g. images) to produce creative visual artifacts.

We propose a system based on Visual Blending and Semantic Network exploration to generate visual representations for introduced concepts (see Fig. 1). The blending process combines existing emoji to create novel ones. The results obtained vary in terms of conceptual complexity, going from literal to metaphoric. We believe that our approach has potential to be explored as an ideation-aiding tool to be used in brainstorming activities, presenting the user with representations for introduced concepts. With this goal in mind, in this paper we value creative and unexpected results and give less importance to literal and unambiguous ones (normally valued in emoji). We present the results of a user study focused on two-word concepts, which analyses the system output in terms of representation quality and surprise degree.

Related Work

Our work addresses two different topics: *Emoji* and *Visual Blending*. As such, we will firstly describe the state of the art for the two topics and then present projects or products which are related to Variation and Customisation of emoji.

Research on Emoji

Previous research on emoji can be mostly divided into the following categories: Meaning, Sentiment, Interpretation, Role in communication, and Similarity between emoji.

Studies on emoji meaning often use word embedding techniques and different data sources (Dimson 2015; Barbieri, Ronzano, and Saggion 2016; Eisner et al. 2016).

In terms of research on emoji sentiment, Novak et al. (2015) provided the first emoji sentiment lexicon, and Hu et al. (2017) compared the sentiments of emoji to the overall sentiment of the message where they occur.

Miller et al. (2016) studied how the users' interpretation of meaning and sentiment of emoji change within and across-platforms, and Rodrigues et al. (2018) addressed how it may differ from intended meanings of developers and researchers.

Some authors address the role of emoji in written communication: Donato and Paggio (2017) studied emoji redundancy and part-of-speech category; Dürscheid and Siever (2017) discussed the function of emoji (complement vs replace); Gustafsson (2017) presented evidence that using emoji to replace words increases reading time; and Wicke (2017) investigated whether emoji could be seen as semantic primes.

Ai (2017) semantically measured emoji similarity. Other authors identified clusters of similar emoji based on emoji vector embeddings (Eisner et al. 2016; Barbieri, Ronzano,

and Saggion 2016). Pohl et al. (2017) used a relatedness-hierarchy to organise emoji. Wijeratne et al. (2017b) created a dataset which contains human-annotated semantic similarity scores assigned to emoji pairs.

On emoji generation, few research work has been conducted and it will be addressed in a later section.

Visual Blending

Visual Blending consists in merging two or more visual representations (e.g. images) to produce new ones. In the context of CC, it is often used together with CB methods to produce representations for a blended mental space. In such cases, it is called Visual Conceptual Blending.

One of the earliest attempts to computationally produce visual blends is, to the best of our knowledge, The Boat-House Visual Blending Experience (Pereira and Cardoso 2002). The work resulted from experiments in interpretation and visualisation of conceptual blends produced for the input spaces *house* and *boat* (Goguen 1999) by an initial version of Divago – one of the first artificial creative systems based on CB theory (Pereira 2007). The visual representations were drawn using a Logo-like programming language.

Ribeiro et al. (2003) used a 3D interpreter to visualise blends of novel creatures produced by Divago from a set of existing ones. The concept maps provided by Divago were converted by the interpreter into Wavefront OBJ files, which could then be rendered.

Steinbrück (2013) presented a framework aimed at exploring the application of CB to the visual domain. It combines image processing techniques with semantic knowledge gathering to produce images in which elements are replaced with similar-shaped ones (e.g. round medical tablets are transformed into globes).

Confalonieri et al. (2015) proposed the use of argumentation to evaluate and iteratively refine the quality of blended computer icons. The authors introduced a semiotic system, which was based on the idea that signs can be combined to convey multiple intended meanings. Despite this, no evidence of a possible implementation was provided.

Xiao and Linkola (2015) presented Vismantic, a semi-automatic system which uses three binary image operations (juxtaposition, replacement and fusion) to produce visual compositions for specific meanings (e.g. *Electricity is green* is represented as the fusion between an image of an electric light bulb with an image of green leaves). The intervention of the user is necessary for both the selection of images and the application of the visual operations.

Correia et al. (2016) developed X-Faces as an approach to Data Augmentation for Face Detection purposes. The system autonomously generates new faces out of existing ones by recombining face parts (e.g. eyes, nose or mouth), using evolutionary algorithms and computer vision techniques.

Cunha et al. (2017) proposed a system for automatic generation of visual blends using a descriptive approach. It used structured representations along with sets of visual relations which describe how the parts – in which the visual representation can be decomposed – relate among each other.

The potential of deep neural networks in tasks related to visual blending has been pointed out by several au-

thors (Berov and Kuhnberger 2016; McCaig, DiPaola, and Gabora 2016; Heath and Ventura 2016). One example is the work DeepStyle (Gatys, Ecker, and Bethge 2015), which explores style transfer in image rendering by recombining the content of an arbitrary image with a given rendering style (e.g. painting styles).

In terms of character blending, one example is the blend of Pokémon (both image and name)[7]. On the same subject, Liapis (2018) produces mappings between type and attributes (e.g. color, shape and in-game sprite), which allow the change of type of a Pokémon.

Current computational approaches to visual blending can be divided into two groups in terms of type of rendering used: the ones which attempt to blend pictures or photorealistic renderings; and the ones that focus on non-photorealistic representations, such as pictograms or icons.

On the other hand, a categorisation can also be done in terms of where the blending process occurs: some interpret or visualise previously produced conceptual blends – e.g. Pereira and Cardoso (2002); others use blending only at the visual level – e.g. Correia et al. (2016); and in others, which can be called hybrid, the blending process starts at the conceptual level and only ends at the visual level – e.g. Cunha et al. (2017).

Variation, Customisation and Generation

Despite the emoji lexicon being constantly increased, there are still a large number of concepts which have not yet found their way into emoji. This is especially evident for more abstract concepts which do not meet the criteria established in the Unicode Guidelines for new emoji. However, several attempts have still been made to complement the system, e.g. *sleep working* by Mentos[8] and *drop the mic* by Microsoft[9]. This shows that the visual representation of more abstract, ambiguous concepts is also valued by the general public.

There are also several examples of user customisation. Windows Live Messenger[10] allowed the user to create emoticons by uploading an image file and Slack[11] currently has the same feature. Some applications allow face-related customisation, e.g. Bitmoji[12], and Taigman, Polyak and Wolf (2016) transform photos of faces into cartoons.

All these examples, serve to show that there is great potential in emoji variation, customisation, and, above all, generation. Despite this, few research work has been conducted on the topic. One example which is related to variation is Barbieri et al. (2017), which investigated the properties of derivations of the kappa emote in Twitch. Specific research on emoji generation mostly uses Generative Adversarial Networks to replicate existing emoji, e.g. (Puyat 2017;

---

[7]pokemon.alexonsager.net, retr. 2018

[8]ementicons.mentos.com/en_GB, retr. 2018

[9]http://huffingtonpost.com/visualnewscom/neil-degrasse-tyson-and-4_b_5615887.html, retr. 2018

[10]news.microsoft.com/2003/06/18/msn-messenger-6-allows-im-lovers-to-express-themselves-with-style/,retr. 2018

[11]get.slack.help/hc/en-us/articles/206870177-Create-custom-emoji, retr. 2018

[12]bitmoji.com, retr. 2018

Figure 3: Visual blends for *rain man* using the same emoji. The first uses juxtaposition and the others use replacement.

Radpour and Bheda 2017). The work of Radpour and Bheda (2017) is particularly interesting, as it is closely related to the idea of our paper by presenting some results for emoji blends. The quality of the results is, however, significantly lower than the one of official emoji, due to visual noise.

The closest work to ours is Emojimoji[13], an emoji generator implemented as part of the Emblemmatic project which also uses Twemoji. It randomly merges emoji shapes and names. However, none of the aforementioned examples uses semantic knowledge in emoji generation, which is the focus of our work.

The Approach

Current needs for more variation and customisation serve as support and inspiration to our main goal: the development of a system that visually represents concepts introduced by the user. This system can be used for several purposes, among which aiding in ideation processes or generating new emoji. Our approach combines data from ConceptNet (Speer and Havasi 2012), EmojiNet (Wijeratne et al. 2017a) and Twitter's Twemoji[14] dataset to explore Visual Blending of emoji.

Resources used

As already mentioned, several resources are put together when developing this system:

– Twitter's Twemoji: a fully scalable vector graphics dataset made available by Twitter. This dataset only consists of images without any semantic information besides the corresponding unicode in the name of each image file. The version used is Twemoji 2.3, which has 2661 emoji;

– EmojiNet: a machine readable sense inventory for emoji built through the aggregation of emoji explanations from multiple sources (Wijeratne et al. 2017a). It was used to provide semantic knowledge to the emoji of the Twemoji dataset despite only having data regarding 2389 emoji;

– ConceptNet: a semantic network originated from the project Open Mind Common Sense (Speer and Havasi 2012). It is used to get concepts related to the one introduced by the user.

The decision to use fully scalable vector graphics is aligned with some of our previous work (Cunha et al. 2017). This image format enables scaling without reducing quality and uses a layered structure – each part of an emoji (e.g. a mouth) is in a separate layer (see Fig. 2). This structure allows an easier blending process and contributes to the overall sense of cohesion among the parts.

---

[13]emblemmatic.org/emojimoji, retr. 2018

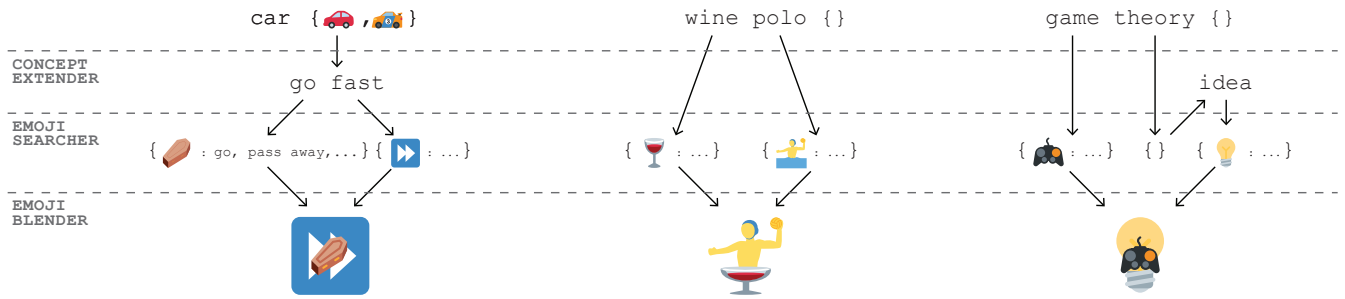[14]github.com/twitter/twemoji, retr. 2018

Figure 4: Generation of visual representations (T2) for three concepts: *car*, *wine polo* and *game theory*

In terms of the semantic knowledge, we initially used the emoji name, emoji definition, emoji keywords and sense definitions – all provided by EmojiNet. However, we concluded that using sense descriptions often leads to unrelated or too specific emoji, which are not useful for the system. For this reason, we decided to use the sense lemmas (word(s) that identify the sense) instead of their descriptions. Unfortunately, the EmojiNet dataset only includes the sense id and its descriptions. In order to solve this problem, the lemmas for each sense id were gathered from BabelNet (Navigli and Ponzetto 2012), which was the original source of the EmojiNet sense data (Wijeratne et al. 2017a).

General Architecture

The system searches existing emoji semantically related to the introduced concept and complements this search with a visual blending process which generates new emoji. In a ideation process, the blending process is useful when there is no existing emoji that matches the concept but also to suggest possible alternatives.

The system consists of two main tasks – retrieval of existing emoji that match the introduced concept (T1) and generation of new ones through visual blending (T2) – which are conducted using three components:

1. Concept Extender (CE): searches ConceptNet for related concepts to the one introduced;

2. Emoji Searcher (ES): searches emoji based on words given, using semantic data provided by EmojiNet;

3. Emoji Blender (EB): receives two emoji as input and returns a list of possible blends.

The system output is a set of visual representations for the introduced concept, composed of existing emoji and generated blends. The system produces a variable number of visual blends, depending on the data found (e.g. Fig. 3).

How it works

The current version works with concepts composed of a maximum of two words. The system starts by analysing the text given by the user. In this first stage, three things can happen: (i) the user introduces a single word (e.g. *car*), (ii) two words (e.g. *wine polo* or *game theory*) or (iii) more. In the last case, the system removes stop-words (e.g. "a", "because", "before", "being", etc.) and considers the result as input text – if after these removal, the word count is still

higher than two, the system ignores it and ends the process without any result.

Retrieval of Existing Emoji (T1) In order to conduct T1, the system mainly makes use of the *Emoji Searcher (ES)* component, which uses EmojiNet dataset to find emoji based on the word(s) given by the user (e.g. in Fig. 4 the *coffin* emoji is retrieved for the word *go* due to its presence in the sense "go, pass away,..."). The word searching is conducted in different places: emoji name and definition, keywords associated with the emoji and senses related to it.

The matching score – i.e. how well an emoji matches the word(s) – is calculated based on the results of the semantic search and the unicode codepoint length ("U+1f474" is more specific than "U+1f474 U+1f3fb"). A value is assigned to each of the criteria:

Name (NV): number of (#) words that match the word(s) searched divided by the total # words in emoji name;

Definition (DV): # words that match the word(s) searched divided by the total # words in emoji definition;

Keywords (KV): (1-1/(# matching keywords))×0.5 + ((# matching keywords)/(total # keywords))×0.5;

Sense (SV): (1-1/(# matching senses))×0.5 + ((# matching senses)/(total # senses))×0.5;

Unicode Codepoint (UV): 1/Codepoint length.

In order to produce the final matching score, the individual values are used together. The criteria have different weights due to importance of each one (e.g. a word in the name is more important than in a sense). Moreover, name, keywords and description were initially gathered from the Unicode Consortium, whereas senses were based on user attribution and may be more ambiguous. The criteria are then weighted according to the following formula: *Emoji matching value* = KV×0.3+NV×0.3+SV×0.2+DV×0.15+UV×0.05

After the searching process is concluded, the system produces a list of emoji that are related to the word given by the user, sorted by *emoji matching value* (e.g. the red and orange cars for the concept *car* in Fig. 4).

Generation of visual representations (T2) In T2 the system behaves differently, depending on the number of introduced words. In the case of single-word concepts, the blending between emoji of the same word does not occur, e.g. two

existing emoji for *car* (the red and orange in Fig. 4) are not blended together to represent the concept *car*. This would only happen if the concept introduced was "car car". Instead, the *Concept Extender* and the *Emoji Searcher* components are used to get the emoji to blend.

The *Concept Extender (CE)* component is used to query ConceptNet for a given word, obtaining related concepts, sorted according to ConceptNet weight system. In the case of single-word introduced concepts, we only consider two-word related concepts (e.g. *go fast* in Fig. 4) as initial experiments indicated that using emoji from two single-word related concepts would result in blends unrelated to the introduced concept. After obtaining the two-word related concepts, the *ES* component (already described for T1) searches for emoji for each word (e.g. in Fig. 4 the *coffin* emoji is obtained for *go*, and the *fast forward* for *fast*). These emoji are then used in the blending process.

On the other hand, when the user introduces a two-word concept, the system firstly searches for existing emoji for each word, using the *ES* component (already described). If emoji are found for both words (e.g. *wine glass* emoji for *wine* and *polo player* for *polo* in Fig. 4), a process of blend is conducted. If the system does not find existing emoji for both words, a search for related concepts is performed, using *CE* component (already described). An example is shown in Fig. 4, in which no emoji is found for *theory*. The system uses the *CE* component to obtain related concepts (e.g. *idea*). After getting the related concepts, the system uses *ES* to search for matching emoji (e.g. *light bulb*). If the search is successful, a blending process is conducted.

The *Emoji Blender (EB)* component is where the blending process occurs, which consists in merging two emoji. The base emoji are selected from the retrieved lists provided by *ES*. In terms of blending, we consider three different methods, even though only two of them are currently being used – these are similar to the ones used in Vismantic (Xiao and Linkola 2015), initially inspired by Phillips and McQuarrie (2004). The first method is *Juxtaposition*, in which the two emoji are put side by side or one over the other (e.g. the blends for *car* and *game theory* in Fig. 4). The second method is *Replacement*, in which part of emoji *A* is replaced by emoji *B* (e.g. in the blend for *wine polo* the water is replaced by wine, see Fig. 4). A blend is produced for each part of emoji *A*: emoji *B* replaces the part using its position (e.g. in Fig. 3 the *rain cloud* emoji replaces the "moustache", the "face shape", the "hair", and the "nose"). The third method is *Fusion*, in which the two emoji are merged together by exchange of individual parts (not used in this paper).

Results and Discussion

In this section we present and discuss the experimental results. We begin by describing an user study and its results. Then, a general analysis of the system and the generated blends is made. Afterwards, we compare the system with previous work, addressing its strengths and shortcomings. In this paper, our goal is to focus on the generation of new visual representations and, for this reason, few attention is given to the process of existing emoji retrieval. In addition,
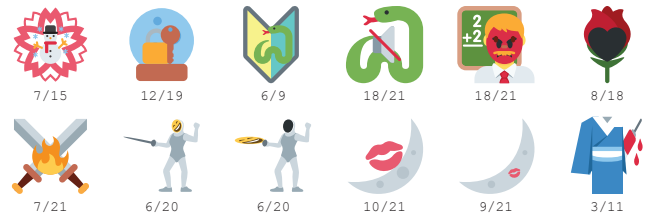

Figure 5: Blends selected as best representation for each concept (top 12). Below each blend is the number of participants who selected it and the total number of participants who selected a blend for that concept. The blends are ordered left-right, top-bottom, according to the order used in Table 1. Two blends are shown for *The Laughing Blade* and *The Sexy Moon*.

we decided to limit our discussion and evaluation to two-word concepts, following our line of research on visual conceptual blending (Cunha et al. 2017). We intend to address single-word concepts in the future.

Evaluating results

In order to assess the quality of system in terms of blend production, a study with 22 participants was conducted. The main goal was to present the participants with blends and ask them to answer a series of questions related to blend quality.

Firstly, a list of ten concepts was produced. These were randomly generated on the website *Title Generator*[15]. The ten concepts are: *Frozen Flower*, *Secrets in the Future*, *Serpent of the Year*, *Silent Snake*, *Storm of the Teacher*, *The Darkest Rose*, *The Flame of the Swords*, *The Laughing Blade*, *The Sexy Moon*, and *The Sharp Silk*. The blends produced by the system for these concepts were shown to the participants. It is important to mention that the number of blends generated is variable and, consequently, the quantity of blends shown was not the same for every concept (e.g. *Silent Snake* has 7 blends and *Storm of the Teacher* has 47).

Each participant saw the blends of every concept but the order in which these were seen was not the same – this was done to minimise the biasing of the results. For each concept, the participants were asked to execute the following tasks: T1 – introduce the concept and generate the blends (presented all at once, side by side); T2 – answer if there is a blend that represents the concept (yes or no); T3 – evaluate quality of representation from 1 (very bad) to 5 (very good); T4 – identify degree of surprise from 1 (very low) to 5 (very high); T5 – select the best blend (only if a positive answer was given to T2). A section for the participants to write optional comments was also included. Asking the user to select the best blend and then make an evaluation of the system based on it may not be the proper way to conduct a user study. However, in the case of our system, it serves the purpose as the end goal is to use it in a process of ideation, in which having at least one good solution is enough.

The results obtained are shown in Tables 1 and 2. Overall, the system was able to generate blends that represented the

---

[15]ruggenberg.nl/titels.html, retr. 2018

Table 1: Number of answers to T2, T3 and T4

Concepts	T2 (represented)		T3 (quality)		T4 (surprise)	
	Yes	No	<4	≥4	<4	≥4
Frozen Flower	13	9	6	7	7	6
Secrets in the Future	18	4	7	11	8	10
Serpent of the Year	5	17	1	4	1	4
Silent Snake	20	2	7	13	7	13
Storm of the Teacher	20	2	5	15	4	16
Darkest Rose	18	4	4	14	10	8
The Flame of the Swords	21	1	3	18	13	8
The Laughing Blade	16	6	12	4	7	9
The Sexy Moon	21	1	5	16	6	15
The Sharp Silk	5	17	4	1	2	3
Total	157	63	54	103	65	92
General Mode				4		4
General Median				4		4

Table 2: Mode and median for T3 and T4 (only includes participants who answered positively to T2)

Concepts	T3 (quality)		T4 (surprise)	
	Mode	Median	Mode	Median
Frozen Flower	4	4	3	3
Secrets in the Future	4	4	4	4
Serpent of the Year	4	4	4	4
Silent Snake	4	4	4	4
Storm of the Teacher	4	4	4	4
The Darkest Rose	4	4	3	3
The Flame of the Swords	4	4	3	3
The Laughing Blade	3	3	3	4
The Sexy Moon	4	4	4	4
The Sharp Silk	3	3	2 and 5	4

concepts – 71.36% (157 out 220) of the answers to T2 were positive (see Table 1) and the quality was above or equal to high (4) in 46.81% (103 out of 220) of the cases.

Moreover, the system is able to produce different blends which can be considered interesting for the same concept. For example, two blends are shown for *The Laughing Blade*, which were selected as the best by the same number of participants (Fig. 5). One reason for this may be the different interpretations for *The Laughing Blade*: a metaphor for the name of the swordsman; or the blade is literally laughing. Similarly, the best blend for *Storm of the Teacher* is metaphoric and for *The Flame of the Swords* is literal. The surprise results seem to reflect this difference: *The Flame of the Swords*, despite having good quality score, was not considered surprising by the majority of the participants, whereas *Storm of the Teacher* was considered both surprising and of good quality.

The worst results were the ones from *The Sharp Silk*, which was only considered concept-representative by 5 participants, from which only one assigned a quality score above or equal to high (4). Their opinion on the surprise criterion was also divided, resulting in two modes (2 and 5).

Most participants reported having difficulty in understanding some of the blends. Some did not recognise a shape

Figure 6: Generation issues. Three blends for *dog* using different related concepts (*drink water*, *guard house*, and *sense danger*, on the left), and blends for *cold* and *unicorn*

(e.g. red shape of *Frozen Flower*), others had different interpretations (a planet instead of a crystal ball for *Secrets in the Future*) and others did not understand the reason behind a blend (e.g. *Serpent of the year*) – see Fig. 5. These were the main reasons for answering negatively to T2, and possibly for the difference in the participants opinion.

General Analysis

Overall, we consider that the results obtained are visually and conceptually interesting (even though no conceptual blending is performed) and, in most cases, unexpected which is supported by the results obtained in the user study.

The system is able to generate variable results, both with the same emoji – e.g. *rain man* in Fig. 3 – and with different ones – e.g. *dog* in Fig. 6. The blending process, through the use of Juxtaposition and Replacement, produces blends that represent the concept behind them and vary in terms of degree of conceptual complexity – in Fig. 1 the blend for *hot dog* is harder to understand than the one for *man bat*. Moreover, the system is able to make less direct connections, e.g. *wine polo* has a literal representation whereas the one for *car* is metaphoric (Fig. 4).

There is no doubt that the performance of the system is dependent on the input emoji and the semantic knowledge associated with it. As such, it might generate interesting blends for some concepts and uninteresting for others. Moreover, in the current implementation, only the emoji with highest matching value is used – changing this would increase the number of resulting visual blends and possibly lead to the generation of better ones (the highest matching value does not necessarily result in the best blends).

The results depend on the word order. The blends generated differ depending on the order of the words introduced. Examples of this are the blends shown in Fig. 1 for *wonder woman* vs *woman wonder* and *apple man* vs *man apple*. Despite already having this in consideration, we think that this connection between name and representation deserves to be further developed, in order to better understand what makes the blend better represent the concept (Pollak et al. 2015).

Results are not always easy to understand. An example of this are the results obtained when introducing the concept *dog* (see Fig. 6). To propose blends for the initial concept, the system makes connections to other concepts. In the case of *dog*, the related concepts are: *drink water*, *guard house* and *sense danger*. Even though all these make sense for describing *dog*, it is not easy to perceive *dog* just by looking at them.

Current issues The blends produced do not always make sense and cannot be considered good representations for the introduced concept– e.g. in Fig. 6 the concept *unicorn* is ex-

tended to *spiral horn*, which then leads to a shell emoji (for *spiral*) and a postal horn emoji (for *horn*). In other cases, the search for related concepts even leads to opposite meanings. This results in the generation of blends that do not represent the introduced concept but something that represents its opposite instead. One example of this is the blend for the concept *cold*, in which a candle is represented (see Fig. 6). Additionally, not all aspects are considered. For example, plurals do not affect the blends in most cases and the removal of stop-words affects the meaning (e.g. *Serpent of the Year* is not the same as *Serpent Year* but the system considers them as equal). These issues make it necessary to further improve the approach in terms of linguistic analysis. However, that was not the focus of this paper and, as such, we do not see these issues as problematic but as future work.

Comparison with previous work

This project can be considered a development of our previous work (Cunha et al. 2017) in the way that both deal with Visual Blending. One major advantage of our approach is that it has a very wide conceptual reach (depending only on the emoji knowledge), whereas in Cunha et. al (2017) the system was limited to the concepts *pig*, *angel* and *cactus*. On the other hand, the present work does not involve Conceptual Blending. We plan on adding semantic information to the initial emoji images, allowing us to implement conceptual blending and thus change the system into a Visual Conceptual Blender.

In comparison to previous research on emoji generation, in which the results distorted by visual noise, we were able to obtain blends of high quality, similar to existing emoji.

Conclusion and future work

We propose a system which has the main goal of generating new emoji by using Visual Blending and Semantic Network exploration. Current state of the art was described, focusing on Emoji and Visual Blending. The architecture of the system was presented and the different system components were explained. In order to assess the quality of the blend generation process, a user study was conducted, which focused on three things: ability to represent concepts, quality of the blends and degree of surprise. Overall the system was able to produce concept-representative emoji and, for many cases, the participants stated that the blends were different from what they were expecting.

Future enhancements to the proposed approach include: (i) increasing the number of words for the concept introduced by the user; (ii) implementing a process of conceptual blending based on Cunha et al. (2017) but also blend evaluation, e.g. (Martins et al. 2015); (iii) defining a fitness function for automatic assessment of blend quality and possibly implementing guided evolution; (iv) cleaning the semantic knowledge from EmojiNet, specially the emoji descriptions which have unuseful information; and (v) exploring blend naming, e.g. (Pollak et al. 2015).

Link The system described in this paper is used on the platform *Emojinating*, which will be available at `http://rebrand.ly/emojinating`.

References

Abbing, R. R.; Pierrot, P.; and Snelting, F. 2017. Modifying the universal. *Executing Practices* 33.

Ai, W.; Lu, X.; Liu, X.; Wang, N.; Huang, G.; and Mei, Q. 2017. Untangling emoji popularity through semantic embeddings. In *ICWSM*, 2–11.

Barbieri, F.; Espinosa-Anke, L.; Ballesteros, M.; Saggion, H.; et al. 2017. Towards the understanding of gaming audiences by modeling twitch emotes. In *Third Workshop on Noisy User-generated Text (W-NUT 2017); 2017 Sep 7; Copenhagen, Denmark. Stroudsburg (PA): ACL; 2017. p. 11-20.* ACL (Association for Computational Linguistics).

Barbieri, F.; Ronzano, F.; and Saggion, H. 2016. What does this emoji mean? a vector space skip-gram model for twitter emojis. In *LREC*.

Berov, L., and Kuhnberger, K.-U. 2016. Visual hallucination for computational creation. In *Proceedings of the Seventh International Conference on Computational Creativity*.

Confalonieri, R.; Corneli, J.; Pease, A.; Plaza, E.; and Schorlemmer, M. 2015. Using argumentation to evaluate concept blends in combinatorial creativity. In *Proceedings of the Sixth International Conference on Computational Creativity*, 174–181.

Correia, J.; Martins, T.; Martins, P.; and Machado, P. 2016. X-faces: The exploit is out there. In *Proceedings of the Seventh International Conference on Computational Creativity*.

Cunha, J. M.; Gonçalves, J.; Martins, P.; Machado, P.; and Cardoso, A. 2017. A pig, an angel and a cactus walk into a blender: A descriptive approach to visual blending. In *Proceedings of the Eighth International Conference on Computational Creativity (ICCC 2017)*.

Danesi, M. 2017. *The semiotics of emoji: The rise of visual language in the age of the internet*. Bloomsbury Publishing.

Dimson, T. 2015. Emojineering part 1: Machine learning for emoji trends.

Donato, G., and Paggio, P. 2017. Investigating redundancy in emoji use: Study on a twitter based corpus. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 118–126.

Dürscheid, C., and Siever, C. M. 2017. Beyond the alphabet–communcataion of emojis. *Kurzfassung eines (auf Deutsch) zur Publikation eingereichten Manuskripts*.

Eisner, B.; Rocktäschel, T.; Augenstein, I.; Bosnjak, M.; and Riedel, S. 2016. emoji2vec: Learning emoji representations from their description. In *Proceedings of The Fourth International Workshop on Natural Language Processing for Social Media*, 48–54.

Fauconnier, G., and Turner, M. 2002. *The Way We Think*. New York: Basic Books.

Gatys, L. A.; Ecker, A. S.; and Bethge, M. 2015. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*.

Goguen, J. 1999. An introduction to algebraic semiotics, with applications to user interface design. In *Lecture Notes in Artificial Intelligence*, volume Computation for Metaphor, Analogy and Agents, 242–291. Springer.

Guibon, G.; Ochs, M.; and Bellot, P. 2016. From emojis to sentiment analysis. In *WACAI 2016*.

Gustafsson, V. 2017. Replacing words with emojis and its effect on reading time. *USCCS 2017* 73.

Heath, D., and Ventura, D. 2016. Before a computer can draw, it must first learn to see. In *Proceedings of the 7th International Conference on Computational Creativity, page to appear*.

Hu, T.; Guo, H.; Sun, H.; Nguyen, T.-v. T.; and Luo, J. 2017. Spice up your chat: The intentions and sentiment effects of using emoji. *arXiv preprint arXiv:1703.02860*.

Lebduska, L. 2014. Emoji, emoji, what for art thou?

Liapis, A. 2018. Recomposing the pokémon color palette. In *Applications of Evolutionary Computation*. Springer.

Martins, P.; Urbancic, T.; Pollak, S.; Lavrac, N.; and Cardoso, A. 2015. The good, the bad, and the aha! blends. In *6th International Conference on Computational Creativity, ICCC 2015*.

McCaig, G.; DiPaola, S.; and Gabora, L. 2016. Deep convolutional networks as models of generalization and blending within visual creativity. In *Proceedings of the Seventh International Conference on Computational Creativity*.

Miller, H.; Thebault-Spieker, J.; Chang, S.; Johnson, I.; Terveen, L.; and Hecht, B. 2016. Blissfully happy or ready to fight: Varying interpretations of emoji. *Proceedings of ICWSM* 2016.

Navigli, R., and Ponzetto, S. P. 2012. Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence* 193:217–250.

Niediek, I. 2016. Don't write it, picture it!: Accessible information by graphic signs. In *Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*, 188–193. ACM.

Novak, P. K.; Smailović, J.; Sluban, B.; and Mozetič, I. 2015. Sentiment of emojis. *PloS one* 10(12):e0144296.

Pereira, F. C., and Cardoso, A. 2002. The boat-house visual blending experience. In *Proceedings of the Symposium for Creativity in Arts and Science of AISB 2002*.

Pereira, F. C. 2007. *Creativity and Artificial Intelligence: A Conceptual Blending Approach*. Berlin: Mouton de Gruyter.

Phillips, B. J., and McQuarrie, E. F. 2004. Beyond visual metaphor: A new typology of visual rhetoric in advertising. *Marketing theory* 4(1-2):113–136.

Pohl, H.; Domin, C.; and Rohs, M. 2017. Beyond just text: Semantic emoji similarity modeling to support expressive communication. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24(1):6.

Pollak, S.; Martins, P.; Cardoso, A.; and Urbancic, T. 2015. Automated blend naming based on human creativity examples. In *23rd International Conference on Case-Based Reasoning (ICCBR 2015) Workshop on Experience and Creativity*.

Puyat, M. 2017. Emotigan: Emoji art using generative adversarial networks. CS229: Machine Learning Course, Stanford University.

Radpour, D., and Bheda, V. 2017. Conditional generative adversarial networks for emoji synthesis with word embedding manipulation. *arXiv preprint arXiv:1712.04421*.

Ribeiro, P.; Pereira, F. C.; Marques, B.; Leitao, B.; and Cardoso, A. 2003. A model for creativity in creature generation. In *4th International Conference on Intelligent Games and Simulation (GAME-ON 2003)*.

Rodrigues, D.; Prada, M.; Gaspar, R.; Garrido, M. V.; and Lopes, D. 2018. Lisbon emoji and emoticon database (leed): norms for emoji and emoticons in seven evaluative dimensions. *Behavior research methods* 50(1):392–405.

Speer, R., and Havasi, C. 2012. Representing general relational knowledge in conceptnet 5. In *LREC*, 3679–3686.

Steinbrück, A. 2013. *Conceptual blending for the visual domain*. Ph.D. Dissertation, Masters thesis, University of Amsterdam.

Taigman, Y.; Polyak, A.; and Wolf, L. 2016. Unsupervised cross-domain image generation. *arXiv preprint arXiv:1611.02200*.

Wicke, P. 2017. Ideograms as semantic primes: Emoji in computational linguistic creativity.

Wijeratne, S.; Balasuriya, L.; Sheth, A.; and Doran, D. 2017a. Emojinet: An open service and api for emoji sense discovery. In *11th International AAAI Conference on Web and Social Media (ICWSM)*, 437–446.

Wijeratne, S.; Balasuriya, L.; Sheth, A. P.; and Doran, D. 2017b. A semantics-based measure of emoji similarity. In *Proceedings of the International Conference on Web Intelligence, Leipzig, Germany, August 23-26, 2017*, 646–653.

Xiao, P., and Linkola, S. 2015. Vismantic: Meaning-making with images. In *Proceedings of the 6th Int. Conference on Computational Creativity, ICCC-15*.