ABSTRACT

A number of empirical studies have demonstrated that certain metaphors can create new conceptual similarities between the source and the target. However, the mechanism of similarity creation has not been explained or modelled satisfactorily. In this paper, we present some empirical evidence in support of this hypothesis for visual metaphors and show how similarity judgements are different from metaphor process. We focus on visual metaphors because here the images corresponding to the source and the target are objectively present. Moreover, we use an algorithmic approach to establish perceptual similarity by using an image-based search system that determines similarity based on low-level perceptual features like colour, shape, texture, etc. Results of this study shows a difference in similarity process and metaphor process, the number of newly created conceptual features also correlated strongly with the perceptual similarity between the pairs of images. Finally, an analysis of the eye-movement data shows that the perceptually similar regions of the images get increased attention during metaphorical interpretation.

INTRODUCTION

It is widely acknowledged that similarities play a key role in interpretation of metaphors. However, there is much debate on what kind of similarities and precisely what role they play in the interpretation process. For example, some researchers distinguish between attribute-based vs. relational similarities. (Gentner & Markman, 1995; Markman & Gentner 1993; Medin, Goldstone & Gentner, 1993) Others, including our own previous work, have focused on the process of representational change to distinguish between before-the-metaphor similarities and after-the-metaphor similarities (Indurkhya, 1992; 1998). We have also analysed metaphor features into 1) source-only features, 2) target-only features, 3) common (to both the source and the target) features, and 4) emergent (novel) features, and have demonstrated that for certain poetic metaphors, many of their features are emergent features. (Gineste, Indurkhya and Scart, 2000).

However, there is little research on how these emergent features come about, and what constrains them (for example, why only certain features emerge and not others.) In our previous work, we had proposed a gestalt interaction view, according to which the conceptual gestalt of the source (or the secondary system) of a metaphor interacts with the perceptual image of the target (or the primary system) to generate these emergent features (Indurkhya, 2006). We have also hypothesized that perceptual similarities (meaning similarities between the perceptual images of the source and the target in terms of perceptual features like shapes, colours, textures, etc.) play a key role in anchoring this process of gestalt interaction (Indurkhya et. al., 2008). In this study we present the results of some of our preliminary experiments to explore this hypothesis empirically.

There are two major issues that need to be resolved in carrying out this ex-
Perceptual vs Conceptual Similarities and Creation of New Features in Visual Metaphors

exploration. One is that the images that the cognitive agents conjure up in their minds as they interpret metaphors are highly subjective, and it is difficult to access them to find out what kind of perceptual features they have. To get around this problem, we chose to focus on visual metaphors, where the images themselves are the stimuli. In other words, in the standard metaphor expression A is B, we make both A and B images for visual metaphor. Though most of the metaphor research has remained focussed on verbal and conceptual metaphors, there have been a few studies on visual metaphors (Danto, 1993, Forceville, 1994, 1996, 2000; Kennedy, 1993; Noel, C, 1994; Rozik, 1994; Simons, 1995; Kaplan, 1990, 1992; Whittock, 1990), and they have been found to exhibit many (but not all) traits of their verbal and conceptual cousins.

The second issue is that given a pair of images, how to determine perceptual similarity between them. If we were to give this task to a cognitive agent, because various conceptual and perceptual processes operate in parallel and interact with each other (Fauconier, G & Turner, M., 1994, 2002), we cannot be sure if the judgment of the agent is based on perceptual similarity alone. To resolve this problem, we employ an algorithmic approach to measure perceptual similarity. In particular, we use a computer image-based search system called FISH (Fast Image Search in Huge Database), (Tandon et. al., 2008) which determines similarity between two images based on low-level perceptual features like color, shapes, texture, etc., to get a perceptual similarity index between a given pair of images.

With this background, we focus on two main questions in this study. One is what kind of features people focus on when making similarity judgment between pairs of images. We hypothesize that it is largely conceptual features. For example, consider the pair of images shown in Figure 1. The image on the left is a well-known Bollywood actress and former miss world Aishwarya Rai. The image on the right is a water buffalo. Now these two images were given a very high perceptual similarity index by the FISH system. In fact, the water buffalo image was retrieved by the system when queried by the Rai image. If you examine them carefully, you can see the perceptual similarities: the light-colored face of the actress framed by dark hair is analogous to the light area between the legs of the buffalo framed by its legs and body. However, when people look at these two images, they tend to focus on conceptual similarities, if they find them similar at all. We further hypothesize that perceptual similarities, even though not consciously noticed, nonetheless influence the creation (or discovery) of conceptual similarities between pairs of images.

The second question concerns the amenability of a pair of images for metaphorical interpretation, and the nature of metaphorical features. Again, we hypothesize that a pair of images with a high perceptual similarity index is more likely to be given a metaphorical interpretation, and that it will also yield a higher proportion of emergent features. For testing this hypothesis, we essentially follow the same methodology as in Gineste, Indurkhya and Scart (2000), where we first gather features of individual images in a feature-generation task, followed by showing the participants visual metaphors in the form given in Figure 2, and asking them for their list of features and interpretations.

Figure 1: Result from FISH

The second question concerns the amenability of a pair of images for metaphorical interpretation, and the nature of metaphorical features. Again, we hypothesize that a pair of images with a high perceptual similarity index is more likely to be given a metaphorical interpretation, and that it will also yield a higher proportion of emergent features. For testing this hypothesis, we essentially follow the same methodology as in Gineste, Indurkhya and Scart (2000), where we first gather features of individual images in a feature-generation task, followed by showing the participants visual metaphors in the form given in Figure 2, and asking them for their list of features and interpretations.
We present here two experiments that shed some light on these questions.

**EXPERIMENT 1 (similarity task)**

**Objective**

To study how [low-level algorithmic] perceptual similarity affects human similarity judgment and the kind of features underlying similarity.

**Participants**

Fifteen (nine females and six males); mean age 22; fluent in English.

**Stimuli**

We used six pairs of images that were given a high perceptual-similarity rating by FISH (*high-similarity pairs*), and six pairs that were given a low perceptual-similarity rating (*low-similarity pairs*).

**Procedure**

Participants were placed in a comfortable chair in front of a Dell 17 inches LCD computer screen placed approximately 45 cms away from their face. They were shown 12 pairs of images, one pair at a time (Figure 2 without “is”), and were asked, “What similarity do you find between the two images in this pair?” They gave their response orally, which was recorded and later transcribed.

**Scoring**

Four subjects were given the list of features and were asked to categorise them into perceptual or conceptual category. To explain these categories, they were given the examples “the rose is red” and “the background is dark” as belonging to perceptual category, and “the rose is beautiful” and “the leaf is alone” as belonging to conceptual category. If they could not say definitely whether the feature belonged to one or the other category, they categorized it as undecided. In (Figure 3) we show a box plot representation of scored features.
Results and discussion:

We found that high-similarity pairs generated more features than low-similarity pairs. The result of the scoring also revealed that for high-similarity pairs, a greater proportion of features were perceptual; whereas for low-similarity pairs a larger number of conceptual features were generated (Figure 4). A one-way analysis of variance revealed that the relation between the similarity index of images and perceptual features was significant: F(1,22)=29.25, p<.001. We did not find any significant effect between perceptual and conceptual features in similarity task.

In one of the low-similarity pairs a ‘leaf’ was paired with a ‘surfer’. We found that participants came up with more conceptual features like ‘both are alone’, ‘both are struggling’, etc. On the other hand, when an image of ‘Taj Mahal’ was paired with an image of ‘wine bottles’, participants came up with more perceptual features like “both of them have similar shape”, and “background is blue in both”.

These results suggest that when there is a high degree of perceptual similarity, it anchors attention of the viewer, and they come up with more perceptual features. But for low-similarity pairs, the viewer is left to connect the images imaginatively, which largely uses conceptual features.

EXPERIMENT 2

Objective

To study how [low-level algorithmic] perceptual similarity affects interpretations of visual metaphors and the proportion of emergent features. In order to associate features with individual images, we first did the following experiment.

EXPERIMENT 2A (Feature extraction)

Participants

Fifteen participants (seven females and eight males); mean age 22; fluent in English.

Stimuli and task

There were 24 images. Participants were asked to describe these images and speak out their features.

Procedure

Images were presented to participants on a Dell 17 inches flat monitor. On seeing the image, the participant was asked to speak out its features. Her or his voice was recorded and later transcribed for further analysis. Once they were done with one image, they pressed a key to go to another image. They could view each image as long as they liked. Also, while they viewed an image their eye movement was recorded with an SR Research Eye link CL eye tracker. (The eye tracker was calibrated to the viewer’s eye at the beginning of the experiment.)

Scoring

The same procedure as in Experiment 1 was used.
Results and analysis

We found that for 24 images participants came up with 315 perceptual features, 165 conceptual features, and 37 undecided ones, as shown in Table 1.

<table>
<thead>
<tr>
<th>Image</th>
<th>Perceptual</th>
<th>Conceptual</th>
<th>Undecided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>12</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Image 2</td>
<td>11</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Image 3</td>
<td>13</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Image 4</td>
<td>10</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Image 5</td>
<td>16</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Image 6</td>
<td>13</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Image 7</td>
<td>15</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Image 8</td>
<td>11</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Image 9</td>
<td>13</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Image 10</td>
<td>13</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Image 11</td>
<td>17</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Image 12</td>
<td>16</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Image 13</td>
<td>11</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Image 14</td>
<td>15</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Image 15</td>
<td>13</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Image 16</td>
<td>17</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Image 17</td>
<td>13</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Image 18</td>
<td>16</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Image 19</td>
<td>11</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Image 20</td>
<td>13</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Image 21</td>
<td>10</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Image 22</td>
<td>13</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Image 23</td>
<td>12</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Image 24</td>
<td>11</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

**EXPERIMENT 2B (Metaphor Task)**

**Participants**

Fifteen participants (Eight females and Seven males); mean age 22; fluent in English

**Stimuli**

We used six pairs of images that were given a high perceptual-similarity rating by FISH (high-similarity pairs), and six pairs that were given a low perceptual-similarity rating (low-similarity pairs).

**Task**

Set up of this experiment was the same as for Experiment 1. Participants were shown pairs of images in the form $X$ is $Y$, where $X$ and $Y$ are two different images (Figure 2), and were asked to describe what this meant to them. If the metaphor is not meaningful, the participants could say so, and proceed to the next pair of images. The responses were given verbally and recorded as in the previous task. Also, the participants’ eye movement was recorded while they carried out this task.

**Scoring**

The same analysis as in Experiment 1 was carried out.
Results and discussion

Figures 5 and 6 show the box plots of features generated during metaphor interpretation for high-similarity pairs and low-similarity pairs. We also found that 87% of the high-similarity image pairs were given some metaphorical interpretation by the participants, as opposed to 41% of the low-similarity pairs. A one-way analysis of variance revealed that the difference was significant at $F(1,28)=39.09$, $p<.001$. We also analysed if the similarity-index of the image had any influence on the metaphor comprehension, and it was found that high-similarity index had a significant effect on the metaphor comprehension $F(1,10)=39.64$, $p<.001$.

We also found a faster response time (mean 3.25 secs.) for high-similarity pairs, as opposed to low-similarity pairs (mean 5.22 secs.) $(F(1,28)=29.59$, $p<.001$).

On further analysis, we confirmed a trend similar to Experiment 1. In particular, more features were generated for high-similarity pairs than for low-similarity pairs. A one-way analysis of variance revealed that features generated in high-similarity pairs ($M=27.33$ SD=1.00) and in low-similarity pairs ($M=11.50$, SD=1.00) differed significantly at the $p<.001$ level $(F(1,12)=19.54$, $p<.001$). Moreover, as in Experiment 1, high-similarity pairs generated relatively more perceptual features, whereas low-similarity pairs generated relatively more conceptual features.

Emergent features

We took all the features generated by the participants while looking at each visual metaphor, and sorted them out into source-only, target-only, common and emergent features as shown in Figure 7 (see also Gineste, Indurkhya and Scart 2000.) Results of this scoring are shown below.

Analysis of eye-movement data:

We analysed the eye-movement data for metaphor task and compared fixation (total time spent in a region of the image) and saccade length (distance between two fixations). It was found that for the high-similarity metaphors, the participants who were able to interpret the metaphor meaningfully had a focused attention on perceptually similar regions of both the images, and a symmetric and uniform saccadic movement between these regions (Figure 8). On the other hand,
in low-similarity pairs we found that though attention was focused on the left image, it was distributed in the right image. Also, average saccade length was smaller in the left image than in the right image.

Figure 8: Heat map and saccade maps in high-similarity pairs.

Figure 9: Heat maps and Saccade map for low-similarity pairs

For further analysis, we considered the main object of the image as the ‘interest area’ leaving out the background. This is shown in Figure 9. We found that number of saccades in the right image of low-similarity pairs (Mean=34) within the interest area was more than the left-image interest area (Mean=21), which suggests that the participants scanned through the images looking for similarity at perceptual level.

For high-similarity metaphors, we also calculated the duration of fixations in the interest area, and found that the mean fixations were almost the same in both interest areas (target: 45; source: 51). This suggests that for generating perceptual similarities, attention is symmetrically divided between the source and the target. In contrast, for low-similarity metaphors, there was a significant difference in the fixation durations (target: 37; source: 63).

Treisman, Sykes, & Gelade, (1977) proposed a new account of attention which assumes that features come first in perception. They assume that the visual scene is initially coded along a number of separable dimensions, such as color, orientation, spatial frequency, brightness, etc. In order to recombine these separate representations and to ensure the correct synthesis of features for each object in a complex display, stimulus locations are processed serially with focal attention. On the other hand, it has also been argued that perception and comparison interact and are mutually dependent (Chalmers et al., 1992). In this view, when a visual stimulus is presented, feature extraction, feature comparison, and conceptualization processes operate in parallel and influence one another.

In our analysis, we assume that a smaller saccade length indicate serial processing. With this assumption, our results suggest a serial processing in all the left images and a tendency of comparison with the right image simultaneously. In high-similarity pairs this phenomenon was found across participants who interpreted images metaphorically. In the case of low-similarity
Perceptual vs Conceptual Similarities and Creation of New Features in Visual Metaphors

pairs, we found serial processing in the left image; but in the right image the saccade length increased. Though it is possible that participants processed features of the right image but could not find a similar perceptual region when they compared it with the left image, and thus we have longer saccades and a distributed attention in the right image.

It was also observed that in low-similarity pairs where participants interpreted images metaphorically, conceptual features of the right image were influenced by the perceptual features of the left image. For example when an image of ‘two red tomatoes’ was shown with an image of ‘a hill,’ participants reported seeing two hills and a small valley in between.

Conclusions and General Discussion

Results of these experiments suggested a difference between the processes of similarity and metaphor. In the similarity task, participants primarily reported existing perceptual similarities between the two images, and it seems that conceptual similarities were considered only when similarity at a perceptual level could not be found. On the other hand, in the metaphor task existing perceptual similarities seem to help in relating the two images conceptually.

Experiment 2 confirms our hypothesis that a pair of images with a high perceptual similarity index is more likely to be given a metaphorical interpretation. It seems that perceptual similarity at the level of color, shape, etc. plays a significant role in metaphor comprehension. This also suggests that our low-level perceptual processes register these similarities and aid the high-level metaphorical process.

We also found that perceptual similarity correlates positively with emergent features that are neither part of the target nor of the source.

An analysis of the eye-movement data shows that when the images are not similar, the left image (target) gets focused attention, whereas the right image (source) gets distributed attention. This might suggest that features are being searched in the source image that might apply to the object of focus in the target image. For high-similarity images, both the target and the source get similar attention patterns, even saccade pattern is symmetrical. This suggests a feature-based comparison. As target is to the left, its features are picked first, which in turn affects the features of the source that are selected.

Another thing we notice is that the source can create new features in the target. Obviously, it happens in low-similarity images as features are searched in the source image that might apply to the target image, and this process can stimulate imaginative ways to render a source-image feature meaningful to the target image. But this can also happen for high-similarity image pairs. For example, when the participants were shown Aishwarya Rai image by itself in Experiment 2, none of them came up with the feature smell. But when this image was paired with the image of a rose (with a high perceptual-similarity index), some participants came up with the feature ‘Aishwarya smells good’.

Future Work

In this study we limited ourselves to conceptually dissimilar images, whether their perceptual similarity index was low or high. So, for example, the two images in Figure 1 are generally judged as dissimilar when the participants are asked to make a quick judgment. Of course, when asked to list out similarity anyway, the participants may generate novel conceptual features to render them similar. In the future we plan to do another study in which we also include conceptually similar pairs of images, with low as well as high perceptual similarity indices.
We should also emphasize that at least for similarity pairs (without the cupola of ‘is’), taking the left image to be the target may be due to the habit of processing information from the left to the right. It would be interesting to see if there are any differences in cultures where the information is normally processed from top to down, or from the right to the left. In our future research, we plan to manipulate the spatial location of images and study its effect on the interpretation and feature generation.

Finally, we would also like to do experiments where the same target image is paired with different source images, and also when the same source image is paired with different target images to better understand the interaction between the two images and to be able to model the mechanism by which conceptual and perceptual features are mapped and new features are created.

REFERENCES

Gentner, D. & Markman A.B. (1995), Similarity is like analogy, in C. Cacciari (Ed.), Similarity, Brussels, Belgium, BRE-POLS
Kaplan, S.J.(1990a), A conceptual analysis of form and content in visual metaphors, Communication.
Kaplan, S.I.,(1990b), Visual metaphors in the representation of communication technology, Critical studies in Mass communication 7, 37-47
Kennedy, J.M., (1993), Drawing and the blind, Pictures to touch, New Haven, CT: Yale University press
Markman, A.B., & Gentner, D. (1993), Splitting the difference: A structural alignment view of similarity, Journal or memory and language 32, 517-535
Medin, D.L., Goldstone, R.L., Gentner, D, (1993),
Respects for similarity, Psychological Review, 100, 254-278

Rozik, E.,(1994), Pictorial metaphors, Kondi
kas/Codes

Simons, J. (1995), Film, Language and conceptual struc-
tual structures: thinking film in the age of cognitivism, PhD thesis, University of Amsterdam, The Netherlands

Tandon, P., Nigam, P., Pudi, V. and Jawahar,

Treisman, A., Sykes, M., & Gelade, G. (1977), Se-
lective attention and stimulus integra-
tion. In S. Dornic (Ed.), Attention and performance VI. Hillsdale, NJ:
Lawrence Erlbaum, pp. 333-36

Whittock, T., (1990), Metaphor and film, Cam-
bridge University press, Cambridge