Semantic Analysis of (Reflectional) Visual Symmetry: A Human-Centred Computational Model for Declarative Explainability

Jakob Suchan, Mehul Bhatt, Srikrishna Varadarajan,
Seyed Ali Amirshahi and Stella Yu

June 2, 2018

Semantic Analysis of (Reflectional) Visual Symmetry

A Human-Centred Computational Model for Declarative Explainability

Jakob Suchan, Mehul Bhatt, Srikrishna Vardarajan, Seyed Ali Amirshahi, Stella Yu

Cognitive Vision. www.cognitive-vision.org

[1]University of Bremen (DE)., [2]Örebro University (SE)., [3]NUST (NO)., [4]UC Berkeley (US)

Abstract

We present a computational framework for the semantic interpretation of symmetry in naturalistic scenes. Key features include a human-centred representation, and a declarative, explainable interpretation model supporting deep semantic question-answering founded on an integration of methods in knowledge representation and computer vision. In the backdrop of the visual arts, we showcase the framework's capability to generate human-centred, queryable, relational structures, also evaluating the framework with an empirical study on the human perception of visual symmetry. Our framework represents and is driven by the application of foundational Vision and KR methods in the psychological and social sciences.

1 INTRODUCTION

The high-level semantic interpretation of symmetry in naturalistic visual stimuli by humans is a multi-layered perceptual phenomena operating at several interconnected cognitive levels. Key aspects include (S1–S4; Fig. 1):

(**S1**). **Spatial Organisation**. high-level conceptual categories identifiable from geometric constructions by way of arbitrary shapes, relative orientation and placement, size of geometric entities, relative distance, depth etc; (**S2**). **Visual Features**. low-level visual features and artefacts emanating directly from color, texture, light & shadow etc; (**S3**). **Semantic Layers**. semantic-spatial layering and grouping based on natural scene characteristics involving, for instance, establishing foreground-background, clustering based on conceptual similarity, relative distance & perceived depth, and application of commonsense knowledge possibly not directly available in the stimulus; (**S4**). **Individual Differences**. grounding of the visual features in the socio-cultural semiotic landscape of the perceiver (i.e., contextual and individualised nuances in perception and sensemaking).

The development of computational cognitive models focussing on a human-centred –*semantic, explainable*– interpretation of visuo-spatial symmetry presents a formidable research challenge demanding an interdisciplinary —*mixed-methods*— approach at the interface of cognitive science, vision & AI, and visual perception focussed human-behavioural research. Our research addresses this interdisciplinarity, with an emphasis on developing integrated KR-and-vision foundations for applications in psychological and social sciences.

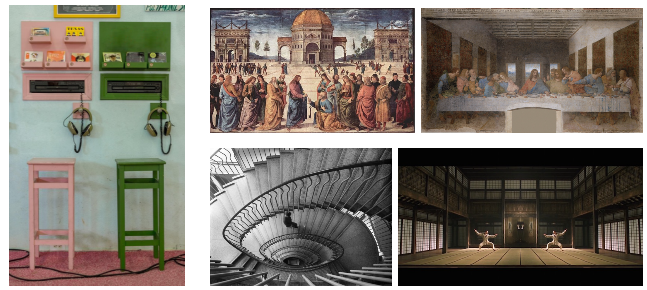


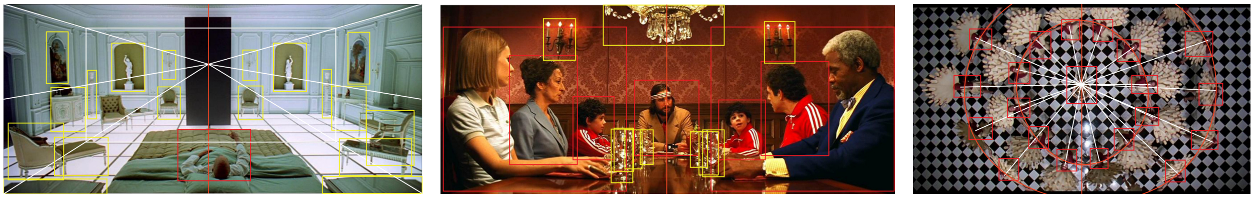
Figure 1: Symmetry perception influenced by visual features, conceptual categories, semantic layering, and nuances of individual differences in perception. Examples include: *"Delivery of the Keys"* (ca.1481) by Perugino, *"The Last Supper"* (1495-98) by Leonardo Da Vinci, *"View of the grand staircase at La Rinascente in Rome, designed by Franco Albini and Franca Helg"* (1962) by Giorgio Casali, and *"The Matrix"* (1999) by The Wachowski Brothers.

Visual Symmetry: Reception – Interpretation – Synthesis

Our research addresses visuo-spatial symmetry in the context of naturalistic stimuli in the domain of visual arts, e.g., film, paintings, and landscape and architectural photography. With a principal focus on developing a human-centred computational model of reflectional symmetry, our approach is motivated and driven by three crucial and mutually synergistic aspects, namely: *reception*, *interpretation*, and *synthesis* (I–III): (**I**). **Reception**. a behavioural study of the human perception (and explanation) of symmetry from the viewpoint of visual attention, and spatio-linguistic and qualitative characterisation(s); (**II**). **Interpretation**. a computational model of deep semantic interpretation of visual symmetry with an emphasis on human-centred explainability and visual sensemaking; (**III**). **Synthesis**. the ability to apply human-centred explainable models as a basis to directly or indirectly engineer visual media vis-a-via their (predictive) receptive effects, i.e., guiding attention by influencing visual fixation patterns (e.g., for marketing), minimising / maximising saccadic movements (e.g., in animation, gaming, built environment planning & design).

Mixed-Methods: Perception – Computer Vision – KR

The semantic interpretation of symmetry requires a mixed empirical-analytical methodology consisting of (M1–M2): (**M1**). **Empirical** / *Human Behaviour Studies*. This involves qualitative studies involving subjective assessments, as well as an evidence-based approach measuring human per-



(a) (b) (c)

Figure 2: Symmetrical structure in visual arts; select scenes from films: a) *"2001: A Space Odyssey"* (1968) by Stanley Kubrick, b) *"The Royal Tenenbaums"* (2001) by Wes Anderson, and 3) *"The Big Lebowski"* (1998) by Joel and Ethan Coen.

formance from the viewpwoint of visual perception using eye-tracking, qualitative evaluations, and think-aloud analysis with human subjects; and (**M2**). **Analytical** / *Interpretation and Saliency*. This involves the development of computational models that serve an interpretation and a predictive function involving: (i). a multi-level computational model of interpreting visuo-spatial symmetry; (ii). a saliency model of visual attention serving a predictive purpose vis-a-vis the visuo-spatial structure of visual media.

With (M1–M2) as the context, this paper focusses on (a–c): (**a**). a computational model principally driven by the capability to generate semantic models that are declaratively grounded and are explainable based on a domain-independent symbolic representation characterising (reflectional) visual symmetry; (**b**). qualitative evaluation with human-subjects, whereby human subjects rank their subjective perception of visual symmetry for a set stimuli using (qualitative) distinctions; (**c**). the technical backbone for assistive technologies for visual media studies (from a psychology viewpoint).

2 SYMMETRY IN ART AND VISION

Symmetry is a well-established stylistic tool applied by visual artists, e.g., painters, photographers, film directors. Symmetry in art is often linked with beauty, and is associated with attributes such as being well-proportioned, well-balanced [Weyl, 1952].[1] Consider the examples from movie scenes in Fig. 2: in the shot from "2001: A Space Odyssey" (Fig. 2a) a centre-perspective is being applied for staging the scene. The symmetry here is obtained by this, as well as by the layout of the room, the placement of the furniture, and the decoration of the room. In particular, the black obelisk in the centre of the frame is emphasising the centre-perspective regularly used by Kubrick, with the bed (and person) being positioned directly on the central axis. Similarly Wes Anderson is staging his shot from "The Royal Tenenbaums" (Fig. 2b) around a central point, but unlike Kubrick's shot, Anderson focuses on the people involved in it. Even though the visual appearance of the characters differs a lot, the spatial arrangement and the semantic similarity of the objects in the shot creates symmetry. Furthermore, the gazing direction of the characters, i.e., people on the right facing left and people on the left facing right, adds to the symmetrical appearance of the shot. In "The Big Lebowski" (Fig. 2c), Joel and Ethan Coen use symmetry to highlight the surreal character of a dream sequence; the shot in Fig. 2c uses radial symmetry composed of a group of dancers, shot from above, moving around the centre of the frame in a circular motion. This is characterised by moving entities along a circular path and centre-point, and the perceptual similarity in the appearance of the dancers.

SYMMETRY AND COMPUTER VISION Symmetry in images has been studied from different perspectives, including visual perception research, neuroscience, cognitive science, arts and aesthetics [Treder, 2010]. Symmetry is an important feature in visual perception and there are numerous studies in vision research investigating how symmetry affects visual perception [Cohen and Zaidi, 2013; Norcia *et al.*, 2002; Machilsen *et al.*, 2009; Bertamini and Makin, 2014], and how it is detected by humans [Wagemans, 1997; Freyd and Tversky, 1984; Csa, 2004]. Most relevant to our work is the research on computational symmetry in the area of computer vision [Liu *et al.*, 2013; Liu *et al.*, 2010]. Typically, computational studies on symmetry in images characterise symmetry as *reflection, translation, and rotation symmetry*; here, reflection symmetry (also referred to as *bilateral* or *mirror symmetry*) has been investigated most extensively. Another direction of research in this area focuses on detecting symmetric structures in objects. In this context [Teo *et al.*, 2015] presents a classifier that detects curved symmetries in 2D images. Similarly, [Lee and Liu, 2012] presented an approach to detect curved glide-reflection symmetry in 2D and 3D images, and [Atadjanov and Lee, 2016] uses appearance of structure features to detect symmetric structures of objects.

► **Computational Analysis of Image Structure** Analysing image structure is a central topic in computer vision research and there are various approaches for different aspects involved in this task. Deep learning with convolutional neural networks (CNNs) provide the basis for analysing images using leaned features, e.g. AlexNets [Krizhevsky *et al.*, 2012] trained on the ImageNet Dataset [Deng *et al.*, 2009]. Most recent developments in object detection involve *RCNN* based detectors such as [Ren *et al.*, 2017; Girshick *et al.*, 2016], where objects are detected based on region proposals extracted from the image, e.g., using selective search [Uijlings *et al.*, 2013] or region proposal networks for predicting object regions. For comparing images, [Zagoruyko and Komodakis, 2015] and [Dosovitskiy and Brox, 2016] measure perceptual similarity based on features learned by a neural network.

3 THE SEMANTICS OF SYMMETRY

Symmetry in visual imagery denotes that an image is invariant to certain types of transformation of the image, e.g. reflectional symmetry is the case where the image does not change,

[1] Symmetry has been employed by artists going back to the masters Giorgione, Titian, Raphael, da Vinci, and continuing till the modern times with Dali and other contemporary artists. (Examples of symmetry are depicted in Fig. 1)

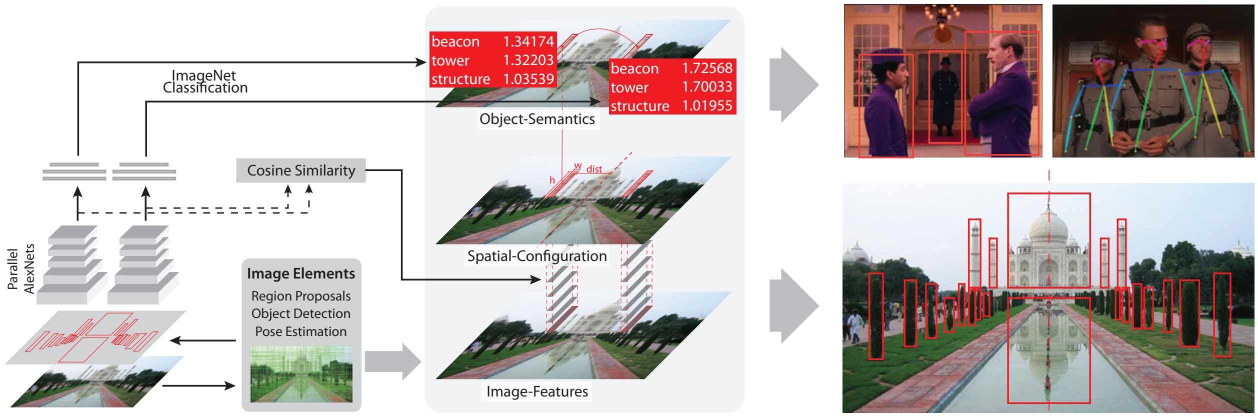


Figure 3: Multi-Level Semantic Symmetry: A Computational Model

when it is mirrored along a specific symmetry-axis. Besides reflectional symmetry, there are various types of symmetry, including rotational symmetry, translational symmetry, etc. Perfect symmetry can be easily detected based on image level features, by comparing pixel in the image, however, in natural images, e.g. coming from visual arts, perfect symmetry is a very rare case and mostly variations of symmetry are used as a stylistic device, where symmetry is only present in some aspects of the image. To address this, we focus on developing a semantic model capable of interpreting symmetrical structures in images.

3.1 A Multi-Level Semantic Characterisation

We develop a multi-level characterisation of symmetry aimed at analysing (reflectional) symmetry. Visual symmetry —in this paper— encompasses three layers (L1–L3; Fig. 3):

L1. *Symmetrical (Spatial) Composition*: spatial arrangement of objects in the scene with respect to a structural representation of a wrt. position, size, orientation, etc.;

L2. *Perceptual Similarity*: perceptual similarity of features in symmetrical image patches, based on the low-level feature based appearance of objects, e.g. colour, shape, patterns, etc.;

L3. *Semantic Similarity*: similarity of semantic categories of the objects in symmetrical image patches, e.g. people, object types, and properties of these objects, such as peoples gazing direction, foreground / background etc.

The proposed characterisation serves as the foundation for analysing and interpreting symmetrical structures in the images, in particular it can be used to identify the elements of the image supporting the symmetrical structure, but also those parts of the image that are not in line with the symmetry, e.g. elements breaking the symmetry, which may be used for investigating the use of balance and in-balance in visual arts, and for analysing how this can be used to guide peoples attention in the context of visual saliency.

3.2 A Model of Reflectional Symmetry

For the computational model presented in this paper (Fig. 3), we focus on reflectional symmetry in the composition of the image based on layers L1–L3 (Sec 3.1), i.e., we investigate image properties based on spatial configuration, low-level feature similarity, and semantic similarity (Alg. 1): we

Algorithm 1: $\mathcal{SYM}(img)$

Data: Image (img)
Result: Symmetrical Structure (SYM)

$\mathcal{R} \leftarrow region\_proposals(img)$
$\mathcal{O} \leftarrow object\_detection(img)$
$\mathcal{P} \leftarrow pose\_estimation(img)$
$\mathcal{E} \leftarrow \{\mathcal{R}, \mathcal{O}, P\}$
for $e \in \mathcal{E}$ **do**
  $f_e \leftarrow AlexNet\_feature\_vec(e)$
for $e_i, e_j \in \mathcal{E}$ **do**
  $COMP_{e_i,e_j} \leftarrow divergence(symmetrical(e_i, e_j))$
  if $COMP_{e_i,e_j} < Thresh$ **then**
    $\mathcal{SP} \leftarrow (e_i, e_j)$
    $SIM_{semantic} \leftarrow conceptual\_similarity(e_i, e_j)$
    $SIM_{percept} \leftarrow cosine\_similarity(f_{e_i}, f_{e_j})$
  $SIM \leftarrow \{SIM_{sem}, SIM_{percept}\}$
$SYM \leftarrow \{\mathcal{E}, \mathcal{SP}, SIM\}$
return SYM

extract object proposals from the image, i.e. regions of the image that are likely to contain an object. These regions are further analysed based on their spatial configuration, the semantics on the objects, and the low-level appearance in terms on image features. We extract elements $\mathcal{E} = \{e_0, ..., e_n\}$ of the image: $(\mathcal{E}_1)$. *Image Patches* are extracted using selective search as described in [Uijlings *et al.*, 2013]; this results in structural parts of the image, e.g., objects, object parts; $(\mathcal{E}_2)$. *People and Objects* are detected in the image using YOLO object detection [Redmon *et al.*, 2016]; $(\mathcal{E}_3)$. *Human Body Pose* body joints and facing direction is extracted using human pose estimation [Cao *et al.*, 2017]. Potential symmetrical structures in the image are defined on the extracted image elements $\mathcal{E}$ by identifying pairs of image elements (symmetry pairs) as well as single elements that are constituting a symmetrical configuration. As such, a potential symmetrical structure consists of image elements e and symmetry pairs $sp = (e_i, e_j)$ where e_i and $e_j \in \mathcal{E}$.

We consider *Compositional structure* (**C1**) of images, and *Similarity* (**C2**) of constituent elements, in particular perceptual similarity in the low-level features, and semantic similarity of objects and regions. The resulting model of symmetrical structure in the image consists of a set of image elements, and the pair-wise similarity relations between the elements.

(**C1**). **COMPOSITIONAL STRUCTURE**. Symmetrical structure is analysed based on the composition of the elements of the image, i.e. using the spatial configuration of the extracted elements and the symmetry axis represented by a vertical line in the centre of the image.

► **Symmetrical Spatial Configuration** Symmetrical composition in the case of reflectional symmetry consists of symmetrically arranged pairs of image elements, where one element is on the left and one is on the right of the symmetry axis, and single centred image elements, which are placed on the symmetry axis. To model this, we represent the extracted image elements as spatial entities, i.e. *points*, *axis-aligned rectangles*, and *line-segments* and define constraints on the spatial configuration of the image elements, using the following *spatial properties* of the spatial entities:

- *position:* the centre-point of a rectangle or position of a point in x, y coordinates;
- *size:* the width and height of a rectangle w, h;
- *aspect ratio:* the ratio r between width and height of a rectangle, i.e., $r = \frac{w}{h}$;
- *distance:* euclidian distance d between two points p and q, i.e., $d = \sqrt{(x_q - x_p)^2 + (y_q - y_p)^2}$.

We use a set of spatial relations holding between the image elements to express their spatial configuration; spatial relations (e.g., $left$, $right$, and on)[2] holding between points and lines describe the relative orientation of image elements with respect to the symmetry axis. Towards this, we use the relative position (rel-pos) of an image element with respect to the symmetry axis, which is defined as the distance to the symmetry axis and the y coordinate of the element.

▷ *Image Patches and Objects* are represented by axis-aligned rectangles, based on the spatial properties of the rectangles, reflectional symmetry is modelled using the following two rules, defining the criteria for a symmetrical configuration of a pair of rectangles, respectively a single rectangles. In the case of a single rectangles e the centre of the rectangle has to be on the symmetry axis.

$$\begin{aligned} &\mathsf{symmetrical}(e) \supset \\ &\quad \mathsf{orientation}(on, \mathsf{position}(e), \mathrm{symmetry\text{-}axis}). \end{aligned} \tag{1}$$

In the case of pairs of rectangles e_i and e_j these have to be on opposite sites of the symmetry axis, and have same size and aspect ratio, further the position of e_i and e_j has to be reflected.

$$\begin{aligned} &\mathsf{symmetrical}(p_i, p_j) \supset \\ &\quad \mathsf{orientation}(left, \mathsf{position}(p_i), \mathrm{symmetry\text{-}axis}) \wedge \\ &\quad \mathsf{orientation}(right, \mathsf{position}(p_j), \mathrm{symmetry\text{-}axis}) \wedge \\ &\quad \mathsf{equal}(\mathsf{aspect\text{-}ratio}(p_i), \mathsf{aspect\text{-}ratio}(p_j)) \wedge \\ &\quad \mathsf{equal}(\mathsf{size}(p_i), \mathsf{size}(p_j)) \wedge \mathsf{equal}(\mathsf{rel\text{-}pos}(p_i), \mathsf{rel\text{-}pos}(p_j)). \end{aligned} \tag{2}$$

The model of symmetry serves as a basis for analysing symmetrical structures and defines the attributes that make the configuration symmetrical.

[2] The semantics of spatial relations is based on specialised polynomial encoding as suggested in [Bhatt *et al.*, 2011] within constraint logic programming (CLP) [Jaffar and Maher, 1994]; CLP is also the framework being used to demonstrate Q/A later in this section.

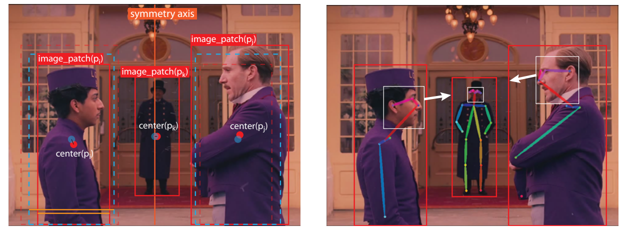


Figure 4: Compositional symmetry: symmetric composition for pairs of image patches, and centering of single image patches

▷ *Human Body Pose* is given by a set of joints j, represented as points, i.e. $pose = \{j_0, ..., j_n\}$. The pose can be either symmetrical within itself, or two people can be arranged in a symmetrical way. Symmetrical body pose is analysed by defining joint pairs $JP = \{(j_k, j_l), ..., (j_m, j_n)\}$, such as $(left\ shoulder, right\ shoulder)$, $(left\ elbow, right\ elbow)$, etc. and compare the relative position of these pairs with respect to the centre of the person c_p.

$$\begin{aligned} &\mathsf{symmetrical}(pose(p)) \supset \\ &\quad \forall (j_k, j_l)\ \mathsf{equal}(\mathsf{rel\text{-}pos}(j_k, c_p), \mathsf{rel\text{-}pos}(j_l, c_p)) \end{aligned} \tag{3}$$

Accordingly, pose of two persons is analysed by defining joint pairs associating each joint of one person to the corresponding joint of the other person, e,g, the left hand of person 1 gets associated to the right hand of person 2.

► **Divergence from Symmetrical Configuration** To account for configurations that are only symmetrical in some aspects, as it typically occurs in naturalistic scenes, we calculate the divergences of the configuration from the symmetry model. For each element of the symmetry structure we calculate the divergence from the defined symmetry model, i.e., we focus on divergence with respect to position, size, aspect ration, and pose (involving configuration of body parts and joints). We use thresholds on the average of these values to identify hypotheses on (a)symmetrical structures.

(**C2**). **SIMILARITY MEASURES**. Visual Symmetry is also based on similarity of image features; we assess similarity of image patches using AlexNets [Krizhevsky *et al.*, 2012] pre-trained on the ImageNet Dataset [Deng *et al.*, 2009], i.e., we use the extracted features to evaluate perceptual similarity and use ImageNet classifications to evaluate semantic similarity of image patches.

► **Perceptual Similarity** Visual Symmetry is based in perceptual similarity of image features, this denotes the similarity in visual appearance of the image patches. To analyse the perceptual similarity of image patches we use the feature vectors obtained from the AlexNets and use cosine similarity to evaluate the similarity of the feature vectors of two image patches. For the case of reflectional symmetry we compare the image patches of all potential symmetry pairs by comparing the features of one image patch to the features of the mirrored second image patch.

► **Semantic Similarity** On the semantical level, we classify the image patches and compare their content for semantic similarities, i.e. we compare conceptual similarity of the predicted categories. Towards this we use the weighted ImageNet classifications for each image patch with WordNet [Miller, 1995], which is used as an underlying structure in Im-

Predicate	Description
sym_element(E)	List of symmetrical elements E.
non_sym_element(E)	List of non-symmetrical elements E.
sym_stats(NP, NSP, MD, MS)	Basic stats on symmetrical structure: number of patches NP, number of symmetrical patches NSP, mean divergence MD, and mean similarity MS.
sym_obj_struct(NO, SO)	Number of symmetrical objects NO and the objects SO.
sym_people_struct(NP, SPPL, SP)	Number of people NP, symmetrical people $SPPL$, and symmetrical pose SP.
non_sym_obj_struct(NSO)	Non-symmetrical objects NSO.
non_sym_people_struct(NSP)	Non-symmetrical people NSP.

Table 1: Sample predicates for querying interpretation model

ageNet, to estimate conceptual similarity of the object classes predicted for the image patches in each symmetry pair. In particular, we use the top five predictions from the AlexNet classifiers and estimate similarity of each pair by calculating the weighted sum of the similarity values for each pair of predicted object categories.

DECLARATIVE SYMMETRY SEMANTICS

The semantic structure of symmetry is described by the model in terms of a set of symmetry pairs and their respective similarity values with respect to the three layers of our model, i.e. for each symmetry pair it provides the similarity measures based on semantic similarity, spatial-arrangement, and low-level perceptual similarity. This results in a declarative model of symmetrical structure, which is used for fine-grained analysis of symmetry features and question-answering about symmetrical configuration in images.

▷ *Extracted Symmetrical Structure*. The Symmetrical structure of an image is given by all elements extracted from the image, the semantic categories, and pairwise similarities.

```
patch(id(0), rectangle(point(233, 53), 107, 466) ).
object(id(0), type(person),rectangle(point(392,106),261,381)).
person(id(0), joint(id(0)), point((582, 159))).
...
category(patch(0), [
        category("file, file cabinet, filing cabinet",
        0.248048), ..., category("desk", 0.166062)]).
...
similarity(perceptual, pair(p636, p1537), 0.429507).
similarity(semantic, pair(p636, p1537), 0.076923).
...
```

Using our framework, it is possible to define high-level rules and execute queries in (constraint) logic programming [Jaffar and Maher, 1994] (e.g., using SWI-Prolog [Wielemaker *et al.*, 2012]) to reason about symmetry. Using the declarative representation of symmetrical structures in images, we can directly *query* symmetrical features of the image, e.g. the following rules characterise symmetrical elements in an image:

```
symmetrical_element(E) :- symmetrical(E).

symmetrical_element(E) :-
        symmetrical(E, _); symmetrical(_, E).
```

Aggregating results for the 'Symmetrical_element(P)' predicate results in (Fig. 6a):

```
SYMETRICAL = [0, 2, 8, 10, 11, 12, 14, 15, 17|...]
```

Similarly we can query for the non symmetrical elements of the image using the following rule:

```
non_symmetrical_element(P) :-
        image_element(P), not(symetrical_element(P)).
```

Resulting in (Fig. 6b):

```
NON_SYMETRICAL = [1, 3, 4, 5, 6, 7, 9, 13, 16|...].
```

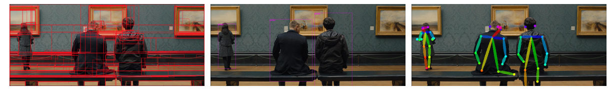
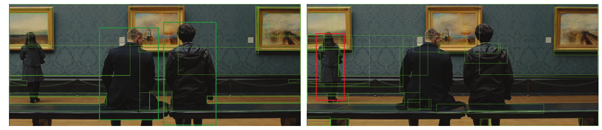


Figure 5: People, objects, regions (*"Skyfall"* (2012) by Sam Mendes).



(a) (b)

Figure 6: Symmetry Analysis: (a) symmetrical elements, and (b) non-symmetrical elements

The underlying interpretation model is queryable using utility predicates (Table 1) in the following manner:

Symmetrical structure of objects and people:

```
?- sym_obj_struct(NumObj, SymObj).

NumObj = 0, SymObj = [].

?- sym_people_struct(NumPeople, SymPeople, SymPose).

NumPeople = 3,
SymPeople = [(person(id(1)), person(id(2)))],
SymPose =   [([upperbody], (person(id(1)), person(id(2)))].
```

Non-Symmetrical structure of people:

```
?- non_sym_people(NonSymPeople).

NonSymPeople = [person(id(0))].
```

Stats on image symmetry:

```
?- sym_stats(NumPatches, NumSymPatches, MeanDiv, MeanSim).
NumImgPatches = 359, NumSymPatches =  40,
MeanDiv = [div_w(12.394), div_h(7.394),
                div_ar(0.944), div_pos(8.32)],
MeanSim = 0.8162167312386968.
```

Hence, our model provides a declaratively interpretable characterisation of reflectional symmetry in visual stimuli.

4 Human Evaluation: A Qualitative Study

EXPERIMENTAL DATASET Human-generated data from subjective, qualitative assessments of symmetry serves many useful purposes: we built a dataset of 150 images consisting of landscape & architectural photography, and movie scenes. The images range from highly symmetric images showing very controlled symmetric patterns to completely non symmetric images. Each participant was shown 50 images selected randomly from the dataset; subjects had to rank the images by selecting one of four categories: not_symmetric, somewhat_symmetric, symmetric, and highly_symmetric. Each image was presented to approx. 100 participants; we calculated the symmetry value as the average of all responses.

EMPIRICAL RESULTS The results from the human experiment suggest, that perception of symmetry varies a lot between subjects. While in the case of no symmetry people tend to agree, i.e. variance in the answers is very low (see Fig. 7), in the case of high symmetry, there is a wider variance in the human perception of symmetry. In particular in the case of images with an average level of symmetry the variance in the answers tends to be high. Qualitatively, there are various aspects on the subjective judgement of symmetry that we can

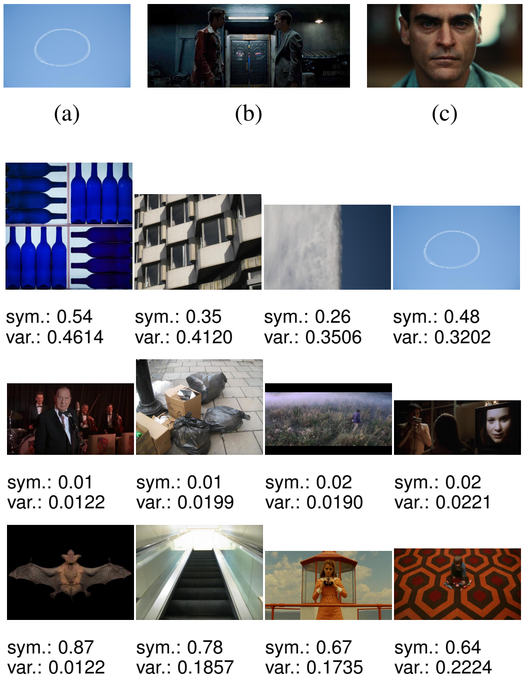


Figure 7: Sample results from the human experiment (Rows 1-3 / bottom-to-top). (**row 1**) most symmetric; (**row 2**) most non-symmetric (these correspond directly to the images with the lowest variance in the answers); (**row 3**) images with the biggest variance in the answers.

Feature Sets	CA (%)	Avg. Sym. Err.	Class Prob. Err.
$fs1$	41.33	0.01806876383	0.0572886659
$fs1$+2	52.00	0.0126452444	0.0400713172
$fs1$+2+3	54.00	0.009900461023	0.0375853705

Table 2: Results from Classification and Prediction Pipeline

observe in the human evaluation (**1 – 3**): (**1**). *absence of features* decreases the subjective rating of symmetry, e.g., the image in Fig. 7a has a nearly perfect symmetry in the image features, but as there are only very few features that can be symmetrical people only perceived it as medium symmetrical, with a high variance in the answers; (**2**). *symmetrical placement of people* in the image has a higher impact on the subjective judgement of symmetry then other objects, e.g. the image in Fig. 7b is judged as symmetrical based on the placement of the characters and the door in the middle, but the objects on the left and right side are not very symmetrical; (**3**). images that are *naturally structured* in a symmetrical way are judged less symmetrical then those arranged in a symmetrical way, e.g. images of centred faces as depicted in Fig. 7c, are rated less symmetrical then other images with similar symmetry on the feature level.

SUBJECTIVE SYMMETRY INTERPRETATION To evaluate how good our symmetry model reflects subjective human criteria for judging symmetry in naturalistic images, we use the results from the human study to train a classifier and a regressor to predict the symmetry class of an image and predict the average symmetry of the images. For our experiment, we extracted three sets of features ($fs1$ - $fs3$) from the symmetry model: $fs1$ consists of the cosine similarity between the two halves of each image on each of the 5 convolution layers in an AlexNet; $fs2$ consists of the symmetrical properties between image patches, i.e., divergence from symmetrical spatial configuration, and perceptual similarity; and $fs3$ consists of the symmetrical properties of object configuration and people in the images. We use the pipeline optimization method of TPOT [Olson *et al.*, 2016] to automatically build the classification and prediction pipelines for the feature sets. The models are trained & tested on the 3-feature set using 5-fold cross validation. Reported are *mean error* & *classification accuracy* (CA) over the 5-folds.

Figure 8: Manual manipulation of symmetry; symmetry decreasing from highly symmetric to not symmetric.

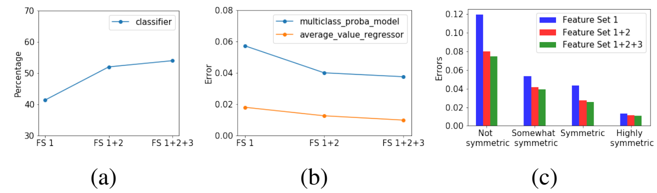


Figure 9: Results of Empirical Evaluation with three different feature set combinations, showing (a) Mean Accuracy, (b) Mean Error, and (c) Class Prob. Error

► *Results and Discussion* The results (Fig. 9; Table 2) show that using the features from our symmetry model improves performance in both tasks, in particular when adding the image patch features $fs2$ we can observe a big improvement in the classification and in the prediction of the average symmetry value (Fig. 9(a), and (b)). Adding people centred features only results in a small improvement, which may be because only a subset of the images in the dataset involves people. The results on the predicted per class probabilities (Fig. 9(c)) shows that by adding features from our symmetry model we are able to much better predict the variances in the human answers, which suggests that the features reflect the human criteria for judging symmetry.

5 OUTLOOK

We have presented a declarative, computational model of reflectional symmetry integrating (visuospatial) composition, feature-level similarity, and semantic similarity in visual stimuli. Immediate next steps focus on: (**1**). *the visual processing aspect*: more advanced region proposals are possible, and can be naturally driven by newer forms of visual computing primitives and similarity measures. The framework is modular and may be extended with improved or new features; (**2**). *space-time aspect*: we go beyond static images to analyse symmetry in *space-time* (e.g., as in the films of Wes Anderson): here, a particular focus is on the influence of space-time symmetry on visual fixations and saccadic eye-movements. (**3**). re-synthesising images (e.g., Fig. 8) to produce qualitatively distinct classes of (a)symmetry, and conducting further empirical studies involving surveys and eye-tracking.

References

[Atadjanov and Lee, 2016] Ibragim R. Atadjanov and Seungkyu Lee. *Reflection Symmetry Detection via Appearance of Structure Descriptor*, pages 3–18. Springer International Publishing, Cham, 2016. 2

[Bertamini and Makin, 2014] Marco Bertamini and Alexis D.J. Makin. Brain activity in response to visual symmetry. *Symmetry*, 6(4):975–996, 2014. 2

[Bhatt *et al.*, 2011] Mehul Bhatt, Jae Hee Lee, and Carl P. L. Schultz. CLP(QS): A declarative spatial reasoning framework. In Max J. Egenhofer, Nicholas A. Giudice, Reinhard Moratz, and Michael F. Worboys, editors, *Spatial Information Theory - 10th International Conference, COSIT 2011, Belfast, ME, USA, September 12-16, 2011. Proceedings*, volume 6899 of *Lecture Notes in Computer Science*, pages 210–230. Springer, 2011. 4

[Bhatt, 2012] Mehul Bhatt. Reasoning about space, actions and change: A paradigm for applications of spatial reasoning. In *Qualitative Spatial Representation and Reasoning: Trends and Future Directions*. IGI Global, USA, 2012.

[Cao *et al.*, 2017] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 1302–1310. IEEE Computer Society, 2017. 3

[Cohen and Zaidi, 2013] Elias H. Cohen and Qasim Zaidi. Symmetry in context: Salience of mirror symmetry in natural patterns. *Journal of Vision*, 13(6):22, 2013. 2

[Csa, 2004] The force of symmetry revisited: symmetry-to-noise ratios regulate (a)symmetry effects. *Acta Psychologica*, 117(3):233 – 250, 2004. 2

[Deng *et al.*, 2009] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009. 2, 4

[Dosovitskiy and Brox, 2016] A. Dosovitskiy and T. Brox. Generating images with perceptual similarity metrics based on deep networks. In *Advances in Neural Information Processing Systems (NIPS)*, 2016. 2

[Freyd and Tversky, 1984] Jennifer Freyd and Barbara Tversky. Force of symmetry in form perception. *The American Journal of Psychology*, 97(1):109–126, 1984. 2

[Girshick *et al.*, 2016] Ross B. Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 38(1):142–158, 2016. 2

[Jaffar and Maher, 1994] Joxan Jaffar and Michael J Maher. Constraint logic programming: A survey. *The journal of logic programming*, 19:503–581, 1994. 4, 5

[Krizhevsky *et al.*, 2012] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012. 2, 4

[Lee and Liu, 2012] S. Lee and Y. Liu. Curved glide-reflection symmetry detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(2):266–278, Feb 2012. 2

[Liu *et al.*, 2010] Yanxi Liu, Hagit Hel-Or, Craig S. Kaplan, and Luc Van Gool. Computational symmetry in computer vision and computer graphics. *Foundations and Trends® in Computer Graphics and Vision*, 5(1–2):1–195, 2010. 2

[Liu *et al.*, 2013] Jingchen Liu, George Slota, Gang Zheng, Zhaohui Wu, Minwoo Park, Seungkyu Lee, Ingmar Rauschert, and Yanxi Liu. Symmetry detection from realworld images competition 2013: Summary and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2013. 2

[Machilsen *et al.*, 2009] Bart Machilsen, Maarten Pauwels, and Johan Wagemans. The role of vertical mirror symmetry in visual shape detection. *Journal of Vision*, 9(12):11, 2009. 2

[Miller, 1995] George A. Miller. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41, November 1995. 4

[Norcia *et al.*, 2002] Anthony M. Norcia, T. Rowan Candy, Mark W. Pettet, Vladimir Y. Vildavski, and Christopher W. Tyler. Temporal dynamics of the human response to symmetry. *Journal of Vision*, 2(2):1, 2002. 2

[Olson *et al.*, 2016] Randal S. Olson, Ryan J. Urbanowicz, Peter C. Andrews, Nicole A. Lavender, La Creis Kidd, and Jason H. Moore. Automating biomedical data science through tree-based pipeline optimization. In Giovanni Squillero and Paolo Burelli, editors, *Applications of Evolutionary Computation*, pages 123–137, Cham, 2016. Springer International Publishing. 6

[Redmon *et al.*, 2016] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 779–788. IEEE Computer Society, 2016. 3

[Ren *et al.*, 2017] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(6):1137–1149, 2017. 2

[Suchan *et al.*, 2016] Jakob Suchan, Mehul Bhatt, and Carl P. L. Schultz. Deeply semantic inductive spatio-temporal learning. In James Cussens and Alessandra Russo, editors, *Proceedings of the 26th International Conference on Inductive Logic Programming (Short papers), London, UK, 2016.*, volume 1865 of *CEUR Workshop Proceedings*, pages 73–80. CEUR-WS.org, 2016.

[Teo *et al.*, 2015] C. L. Teo, C. Fermüller, and Y. Aloimonos. Detection and segmentation of 2d curved reflection symmetric structures. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1644–1652, Dec 2015. 2

[Treder, 2010] Matthias Sebastian Treder. Behind the looking-glass: A review on human symmetry perception. *Symmetry*, 2(3):1510–1543, 2010. 2

[Uijlings *et al.*, 2013] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders. Selective search for object recognition. *International Journal of Computer Vision*, 104(2):154–171, 2013. 2, 3

[Wagemans, 1997] Johan Wagemans. Characteristics and models of human symmetry detection. *Trends in Cognitive Sciences*, 1(9):346–352, 1997. 2

[Weyl, 1952] H. Weyl. *Symmetry*. Princeton paperbacks. Princeton University Press, 1952. 2

[Wielemaker *et al.*, 2012] Jan Wielemaker, Tom Schrijvers, Markus Triska, and Torbjörn Lager. SWI-Prolog. *Theory and Practice of Logic Programming*, 12(1-2):67–96, 2012. 5

[Zagoruyko and Komodakis, 2015] Sergey Zagoruyko and Nikos Komodakis. Learning to compare image patches via convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 4353–4361. IEEE Computer Society, 2015. 2