Publication information:
Frontiers of Computer Science, Vol. 10, No. 2, pp. 216-232, 2016.
Front. Comput. Sci.
DOI

Cognitive mechanism related to line drawings and its applications in intelligent process of visual media: A Survey

Yongjin LIU(✉)[1], Minjing YU[1], Qiufang FU[2], Wenfeng CHEN[2], Ye LIU[2], Lexing XIE[3]

1 Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
2 State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
3 Research School of Computer Science, The Australian National University, Australia

Abstract Line drawings, as a concise form, can be recognized by infants and even chimpanzees. Recently, how the visual system processes line-drawings attracts more and more attention from psychology, cognitive science and computer science. The neuroscientific studies revealed that line drawings generate similar neural actions as color photographs, which give insights on how to efficiently process big media data. In this paper, we present a comprehensive survey on line drawing studies, including cognitive mechanism of visual perception, computational models in computer vision and intelligent process in diverse media applications. Major debates, challenges and solutions that have been addressed over the years are discussed. Finally some of the ensuing challenges in line drawing studies are outlined.

Keywords line drawings, cognitive computation, visual media, intelligent process

1 Introduction

Visual media usually includes images, videos and 3D digital models. Intelligent process of visual media plays an important role in many areas of academic research and industrial applications. Nowadays, with the development of digital media capturing devices and internet techniques, big media data is emerging. How to efficiently process, obtain required information and reuse big media data is a currently hot research area in computer science.

Human brain has remarkable capacity for visual media processing. In [1] there is a good comparison between human brain and the state-of-the-art supercomputer: 1) in human brain, memory is 3.5 quadrillion bytes, computing performance is 2.2 billion megaflops and the power is 20 watts, 2) in the world's most powerful supercomputer in 2011, the K from Fujitsu, memory is 30 quadrillion bytes, computing performance is 8.2 billion megaflops and the power is 9.9 million watts. A conclusion was drawn in [1] that “computers are good at storage and speed, but brains maintain the efficiency lead’’. Accordingly, cognitive science that is the interdisciplinary scientific study of the human mind and its processes has attracted considerable attention recently.

Recently, line drawings had been emerging as an efficient tool for intelligent processing of big visual medial data. A line drawing (sometimes also called a sketch) is a set of sparse, simple two- dimensional feature lines without hatching or stippling for shading/tone effects. Fig.1 shows an example, in which a line drawing (Fig.1b) is extracted from a color photograph (Fig.1a), by a trained artist. Obviously, the line drawing only carries a fraction of original information such as lines and shapes, and the surface information such as color and texture is lost. However, human can easily understand the line drawings at a glance. In this survey paper, we first introduce the cognitive mechanism related to line drawings in Section 2. Then in Section 3, we summarize several popular computational models of visual perception for line drawings. The intelligent process and applications of line drawings are presented in Section 4. Section 5 presents several future directions and finally Section 6 concludes this paper.

(a) (b)

Fig.1 A line drawing (b) produced by a trained artist at the Academy of Arts & Design, Tsinghua University, by selectively tracing contours in a color photograph (a)

2. Cognitive mechanism of line-drawing perception

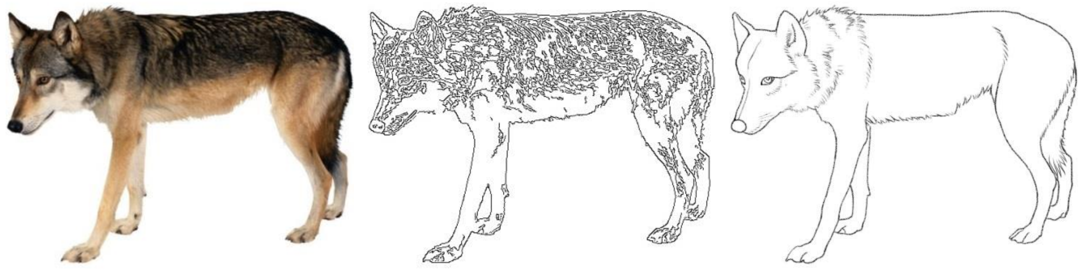
Based on recent findings in neurophysiology and cognitive psychology, the cognitive process of visual media is divided into three successive stages in [2], namely perception, memory and judgment (PMJ). The perception stage concerns how sensory information is perceived. The memory stage concerns how to encode and store conceptual prototypes as well as how to update and consolidate them for both short-term and long-term memories. The judgment stage concerns how to make a decision to complex problem solving.

One of the fundamental problems in vision is visual recognition [3], in which two related but different topics are object recognition and natural scene categorization. A distributed computational cognitive model for object recognition using the PMJ model was studied in [4]. One object recognition model in [5] provided evidence for partitioning the recognition process into PMJ stages: 1) Processing of basic object components, such as form and color. These basic components are then grouped to provide information on distinct edges for a visual form, and subsequently lead to figure-ground segregation. 2) The visual representation is matched with structural descriptions in memory. 3) Semantic attributes are applied to the visual representation, providing meaningful recognition.

Below we summarize the cognitive mechanism related to line drawings in object recognition and natural scene categorization.

2.1 Surface versus edge-based representation

Much information can be used for object or scene recognition, such as shape, color, brightness, texture and motion cues, etc. These information can be classified into two classes:



(a) Original color photograph (b) Canny edge extraction (c) A line drawing by an artist

Fig.2 Surface vs. edge-based representation. (a) The original color photograph contains surface gradient. (b) Edge extraction using the Canny's method [6]. (c) A sparse line drawing from a trained artist by selectively tracing contours in the color photograph

1) edge-based information, including the shape information represented by contours or line drawings, 2) surface-based information, including color, brightness and texture, etc.

Edge-based information is mainly contours or lines that intend to convey specific 3D shapes. Not every edge corresponding to a sharp gradient change in a color photograph is used in a line drawing. Indeed, a line drawing favored by the visual system is only a subset of these edges and it seems that only artists capture these ambiguous rules [7]. See Fig.2 for an example. These sparse line drawings only consist of major components of an object, while small details related to texture, shading and shadows are usually discarded. Line drawings seem to be appropriate for the visual system because human brain has remarkable capacity to deal with line drawings [8], and even infants [9] and chimpanzees [10] can recognize line drawings. For example, a child can easily recognize the curved shape from the sparse line drawing shown in Fig.3 as a hat. An interesting study investigated how well line drawings depict shape [11] based on a gauge figure protocol in visual psychophysics [12, 13]. Their findings revealed that people interpret certain shapes from a line drawing almost as well as those from a color photograph.

(a) (b)

Fig.3 A comparison of surface-based (a) and edge-based (b) representations: a person can easily perceives the curved shape of the hat from a sparse line drawing (b), even without fine details (i.e., surface information such as texture, color, shading and shadows) in the original color photograph (a)

Surface-based information is useful to infer surface material property and 3D shape. For example, from variations in color or brightness, a 2.5D depth estimation called 2.5D sketch can be constructed [14]. Watt and Rogers [15] argued that 2.5D is the construction of a three-dimensional environment from 2D retinal projections. The surface-based information towards a 2.5D sketch representation is also well studied in computer vision, in which the so-called shape from X technique (X stands for shading, texture, motion, focus/defocus, etc) is still a hot research area [16]. The surface material property inherent in surface-based information also provides additional distinct features for facilitating object recognition. For example, for the high color diagnostic (HCD) objects such as banana, the detection of color such as yellow will facilitate the recognition. Similarly, textures such as those in zebra are also helpful for the recognition.

Although both edge-based and surface-based information make a contribution to object recognition, there is a long debate on how they contribute to the recognition. An edge-based theory assumes that edge-based information is sufficient for object recognition and surface gradients provide less efficient routes for accessing the memorial representation [17, 18]. For example, it is found that reaction times and error rates were virtually identical for common objects of color photographs and line-drawings when the images were briefly (50-100 ms) presented. That is, the edge-based representation is sufficient for object recognition and as stated in [18], "although differences in surface characteristics such as color, brightness, and texture can be instrumental in defining edges and can provide cues for visual search, they play only a secondary role in the real-time recognition of an intact object when its edges can be readily extracted."

On the contrast, an alternative surface-based theory assumes that surface gradients are central for object recognition and both edge-based and surface-based information are simultaneous routes for basic level categorization. For example, it is found that color improved object recognition of common food items when there was no time limit in stimulus presentation. Notice that the stimulus duration might lead to the different findings. We [19] manipulated the *Stimulus Onset Asynchrony* (SOA) to examine this and found that the stimulus duration did mediate the role of surface gradients in natural scene categorization. When the SOA was short enough, the accuracy was higher for line-drawings than for color photographs, although when the SOA was longer, the accuracy was lower for line-drawings than for color photographs. The findings were consistent with the edge-based theory, providing new evidence for contour information receiving priority processing.

Moreover, it is widely believed that the visual system operates in a way of performing a strong data reduction at an early stage [20], and creates a compact summary of relevant information, determining what is perceived as "meaningful features", that can be handled by further levels of processing. The initial processing of visual information is often described as the extraction of a simplified line drawing based on a limited number of ‘‘salient features" [21].

2.2 Features in the recognition

The performance of object recognition and natural scene categorization in human vision system is subject to the image features, such as shape, color, luminance, contrast, orientation, texture, Fourier spectra (amplitude and phase spectra) and emotional meaning. The relations between performance and these individual features have received much investigation. See [22,23] for some comprehensive summaries.

Below we summarize some widely studied features in vision research.

- A shape is usually represented by external boundary or outline of an object, detected in images by difference between the object and the background.
- Color is a visual perceptual property in human brain characterized by the color categories such as red, green, blue and others. The color of an object depends on 1) the physics of the object, 2) its embedded environment and 3) the

characteristics of the perceiving eye/brain (or camera).

- In optical science, luminance is a photometric measure about the amount of visible light leaving a point on a physical surface or an imaginary plane in a given direction. In a gray scale image, the luminance value is usually the pixel value. In a color image, the luminance value is usually determined by some weighted combinations of three color channels, e.g., OpenCV uses grayscale = 0.299×red+0.587×green+0.114×blue. In [23], a contrast in color between the animal and its background is defined physiologically in a DKL color space [24].
- In vision, contrast is the difference in luminance and/or color between parts of an image. The human visual system is more sensitive to contrast than absolute luminance. Several definitions exist. A contrast definition used in physiology [25] is the maximum luminance minus the minimum luminance divided by twice the mean luminance. Some other definitions such as Weber contrast, Michelson contrast and root mean square contrast, are also used.
- Orientation is a placement of an object in a rotational coordinate system with respect to a fixed point and a reference position.
- Texture is a repetitive pattern with a specific luminance, contrast, color and orientation, and thus can be regarded as a multi-cue feature. In other words, texture is not independent from its constituting features.
- The Fourier transform maps signals such as images from spatial domain to frequency domain. In the frequency domain, a discrete set of complex amplitudes, called Fourier series coefficients, represent the frequency spectrum of the original signal. The complex number gives both the amplitude (or size) of the wave and the phase (or the initial angle) of the wave.

In Sections 2.1 and 2.3, we summarize neuroscientific studies based on the comparison of line drawings (shape feature) and color photographs (color feature) in natural scene categorization. These works can also be extended to consider more features, e.g., a set of grayscale photographs of natural scenes equalized in average luminance, global contrast and spectral energy were used in a phase feature analysis in rapid visual categorization [26].

In [27], a series of experiments revealed systematic effects of orientation on the time required to identify line drawings of natural objects. The result suggested that novel depictions of a known class of objects may be identified by a process of mental rotation. That is, to recognize misoriented objects, our visual system may normalize the stimulus representation to a canonical upright position through a process of mental rotation and then recognize the upright image. However, this hypothesis was skepticized in [28] by observing a patient with a large right middle cerebral artery territory stroke who fails three different mental rotation tasks but is nonetheless able to recognize misoriented numbers, letters, and drawings.

For the role of amplitude spectrum and phase alignment, some researchers [25, 29] regarded that the visual system behaves as a Fourier analyzer and the Fourier components of an image are represented as amplitude and phase spectra. It was a common belief that phase spectrum can determine most recognizable image structure [30] and phase information dominates the perception of natural scenes [31]. However, some recent studies showed that amplitude-based processes

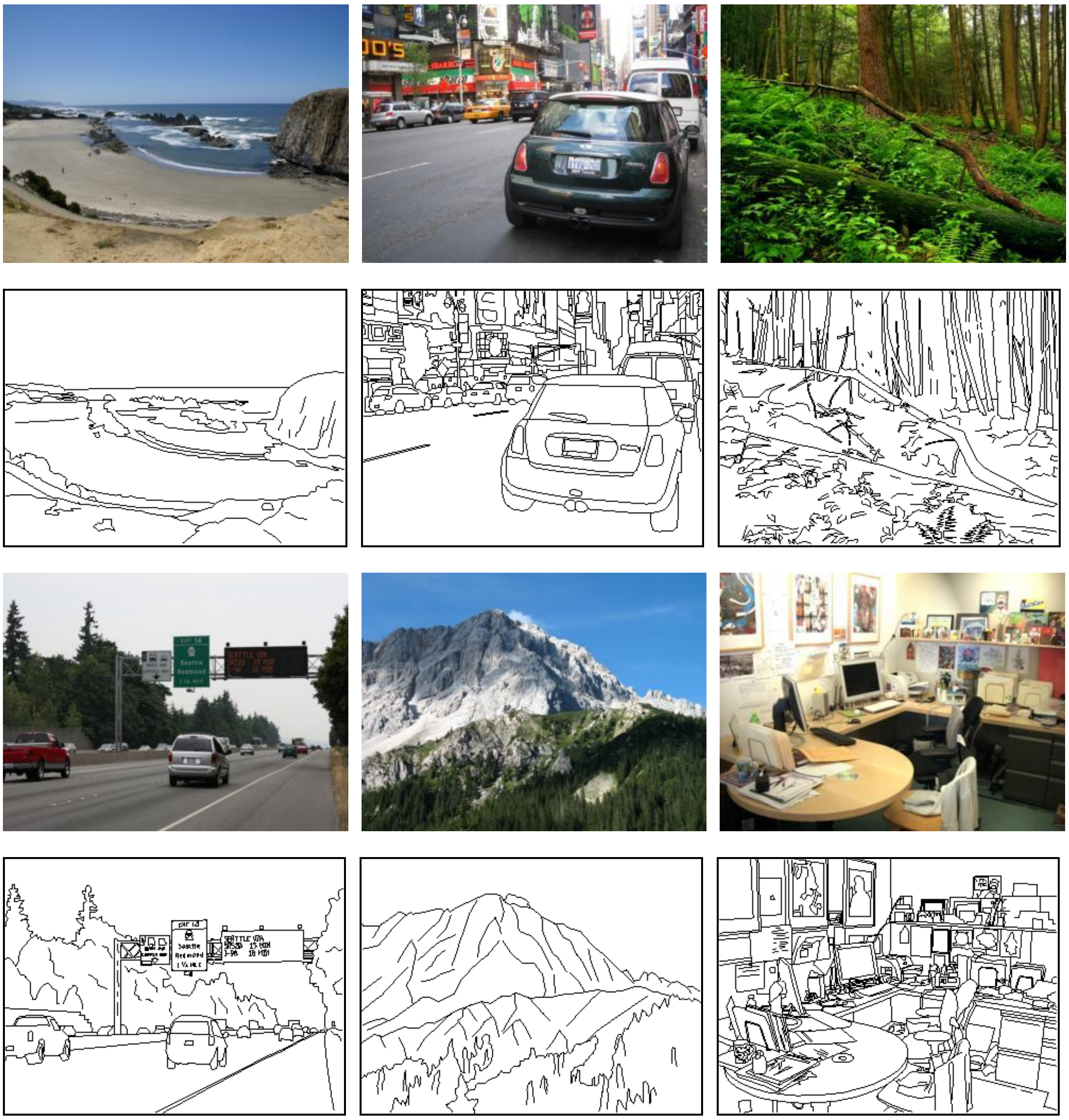


Fig.4 Photographs and line drawings of six categories of natural scenes: beaches, city streets, forests, highways, mountains and offices. Original images by flickr users David K, Nicholas A. Tonelli, francois, Norris Wong, David Herrera, and Mo Riza

are sufficient for rapid scene categorization [32] and amplitude spectrum characteristics of the natural scenes are useful to speed up context categorization processes [26]. Our study in a subliminal perception experiment showed that compared to the gray photographs, line drawings is less influenced by amplitude spectra [33]. This is in line with [31] that amplitude information was less important than phase for perception, and lines and edges may be mainly defined by phase information, providing a useful definition of visual features [34].

2.3 Patterns of brain activity measured by fMRI

Line drawings (mainly shape features) and color photographs (multi-cues features) have remarkable difference in image statistics (Fig. 4). To reveal how human brain processes line drawings and color photographs, the functional magnetic resonance imaging (fMRI) technique was widely used.

The fMRI measures brain activity by detecting the changes in blood oxygenation and flow that occur in response to neural activity [35]: a more active brain area consumes more oxygen for meeting an increased demand of blood flow, and activation maps can be produced to show which parts of the brain are involved in a particular mental process.

The fMRI studies showed that a set of lines that match the cube's edges would trigger similar occipital responses as the original cube indicating that, on a neural level, line representations are equivalent to the originals they depict [8]. Importantly, a recent fMRI study [36] showed that the areas in primal visual cortex, such as V1, the parahippocampal place (PPA), retrosplenial cortex (RSC), and lateral occipital complex (LOC), were responsible for extracting differences in spatial layout among different scene categories. Based on this finding, a further fMRI study [37] showed that in the PPA and RSC areas, the line drawings generate similar neural activation as color photographs, indicating that human vision system may use a schematic representation analogous to simple line drawings for encoding and processing scene category information. Moreover, by selectively removing long or short contours from the line drawings, they [37] found that the global scene structure preserved in line drawings (possibly corresponding to low frequencies in the amplitude spectrum) plays a crucial role in representing scene categories.

Although the fMRI studies provided neuroscientific evidence that edge-based representation akin to simple line drawings are sufficient for rapid object and scene category recognition, it remains unclearly whether the edge-based information receives priority processing because of the poor temporal resolution of the fMRI studies on the order of one or few seconds. To determine the time course of object and scene recognition, the event-related potential (ERP) technique was used [38].

2.4 Different time courses in visual perception

Electroencephalography (EEG) is a non-invasive and inexpensive method using electrodes placed on the scalp to record the electrical activity of large populations of neurons that are synchronously firing in the brain over time. By means of EEG, an event-related potential (ERP) experiment can be designed, in which a large number of time-locked experimental trials are averaged together, causing random brain activity to be averaged out and the relevant waveform to remain. The ERP technique can investigate the perceptual process with millisecond precision. This high temporal resolution makes available the study of transitions from earliest sensory-based perceptual processes to the higher cognitive processes.

As suggested in [39, 40], object recognition consists of two consecutive stages: an early perceptual stage about 75 milliseconds and a later stage for decision making about 150 milliseconds. To address the differences in the time course of categorizing color photographs and line drawings, an ERP study [41] was conducted by using a categorization task with backward masking technique, where a variable delay between the scene and the mask (the Stimulus Onset Asynchrony, SOA) was manipulated. The ERP results showed that overall, the latencies of early ERP components were slower for color photographs than for line-drawings, and the latencies of early ERP components significantly increased with longer SOA only for color photographs but not for line-drawings. These results suggest that more usable information continues to be extracted from color photographs rather than line drawings as SOA increased and provided strong evidence for the edge-based theory.

2.5 Semantic access to conceptual representation in human brain

An old Chinese saying is that "a good picture is worth a thousand words." In our work, we conjecture that a good simple line drawing, akin to Hieroglyphics, behaves like a conceptual prototype that provides a semantic access to conceptual representation in our brain [4, 42]. Therefore line drawings play an important role in bridging the big gap between low-level image feature processing and high-level semantic understanding. Accordingly, line drawings are considered to be the simplest and most typical pictorial representations of objects and scenes. Below we summarize the applications of line drawings in psycholinguistic and cognitive research.

Snodgrass and Vanderwart [43] published their classic corpus of 260 black-and-white line drawings of objects from 14 concrete categories. As the first standardized set of pictures, all the line drawings were standardized on four variables of central relevance to memory and cognitive processing, that is, name agreement, image agreement, familiarity, and visual complexity [43]. During the last three decades, this historical article [43] received more than 3100 citations. This corpus has been extensively used in experimental and clinical research on cognitive processing, such as language, memory, and object recognition [44]. After Snodgrass and Vanderwart's seminal work, several new test sets of picture norms were published, e.g., in [44-46], most of which were comprised of line drawings.

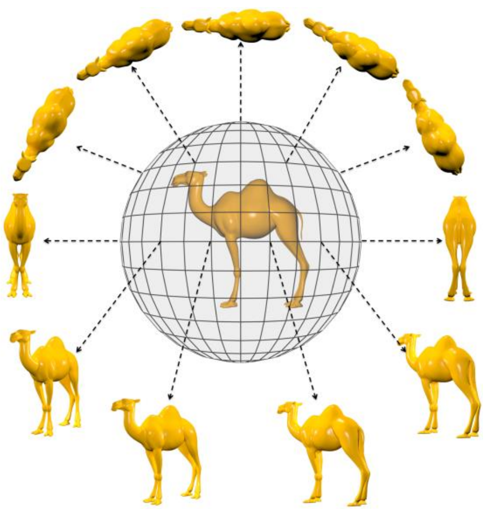
In the studies of language production, line drawings presented in [43] were used as stimuli in language production tasks to investigate the activation of conceptual representation and the following lexical access, e.g., in [47-50]. For all of language production models, their ultimate aim is only one, which is to describe all of the stages between having a concept and translating that concept into linguistic form. The general paradigm of this area to make participants in the tasks for producing language is to present them a line drawing of an object, and this paradigm has been proved a very efficient way to address the questions of language production so far. It seems that line drawings of objects could be perfect visual symbols for semantic access to their conceptual representations in human brain.

3. Computational models of line drawings

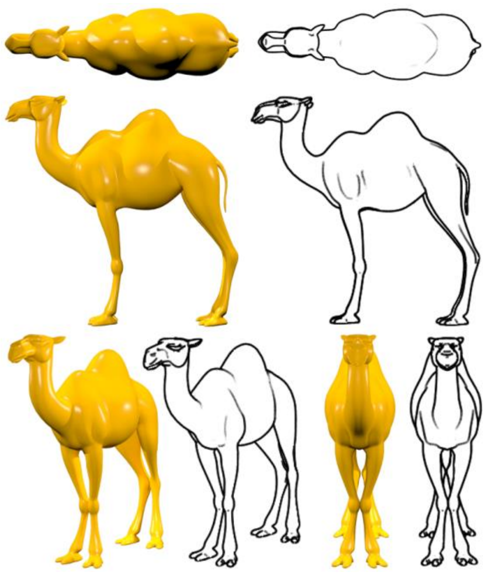
As revealed in neurophysiological evidence summarized in Section 2, there are large amount of redundancies in natural photographs and line drawings can serve as a concise form for fast processing of visual media. To utilize these neurophysiological findings, in this section, we summarize computational models inherent in line drawings' analysis and interpretation.

We use the PMJ model [2] to characterize the cognitive process involving line drawings by the following perception, memory, learning and judgment stages.

- *Perception starting with image projection*. Human perceives the world through eyes and images are projected onto the retina. We simulate this process by sampling viewpoints from a spherical distribution (Fig. 5a). The projected images are then converted into line drawings for a fast process in later stages (Fig. 5b).
- *Memory coding as local descriptors by exploring geometric structures*. We



(a)



(b)

Fig.5 (a) Image projection using viewpoints uniformly sampled from a spherical surface and (b) the images are converted into line drawings using the CLD method [51]

simulate the short-term memory for storing local abstract forms by exploring geometric structures inherent in the line drawing (Fig. 6).

- *Learning in feature space*. The local abstract forms encode local geometric structures of line drawings in a feature space and these abstract forms are evolving in time for different user experiences, served as the long-term memory.
- *Judgment using local descriptors*. Most judgment tasks can be efficiently achieved by using the vocabulary in the long-term memory.

3.1 Perception starting with image projection

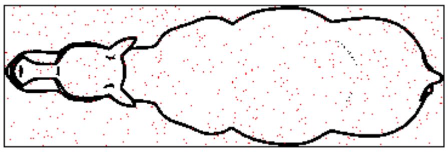
Human views the physical world through the images projected on retina. Many methods have been proposed to determine the representative images of 3D objects. Broadly there are two classes. One is the good viewpoint selection and the other is to cluster uniformly sampled viewpoints into representative viewpoints.

For the first class of good viewpoint selection, the information associated with different viewpoints is measured by some quantitative criteria and few good viewpoints are determined by solving an objective function (e.g., [52, 53]). An entropy map method is proposed in [52], in which the Shannon entropy is used as a measure. An image x may best represent one of *n* objects $\{O_i\}$. The characteristic of each representation can be modeled by a posterior probability distribution P(O|x) and its Shannon entropy is $\mathrm{H}(\mathrm{P}(\mathrm{O}|\mathrm{X})) = \sum_i P(O_i|x) log \frac{1}{P(O_i|x)}$. The measure $p(O_i|x)$ is then determined by standard Bayesian method $p(O_i|x) = \frac{1}{K} p(m_x|O_i) p(O_i)$,

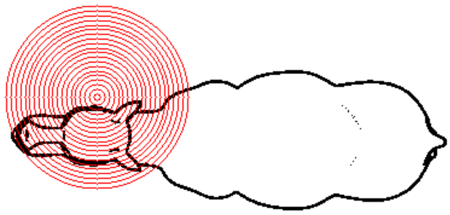
where K is a normalization constant and m_x is determined in a training preprocess. Recent work on good viewpoint selection makes use of human perception with the observation that the human-preferred view is highly correlated with a combination of several simple measures such as silhouette length and projected area [54].

In the second class, an object is placed at the center of a bounding sphere (Fig5a). Then

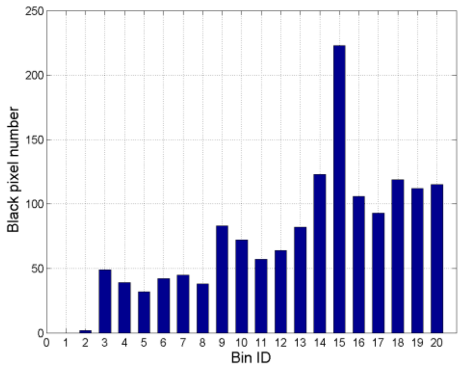
(a)



(b)



(c)



(d)

Fig.6 (a) The minimum-area enclosing box B. (b) Randomly sample 500 points in B. (c) a circular region C centered at a sample point, which partitions into 20 circular rings. (d) The circular histogram of (c).

a set of dense and uniform viewpoints are sampled on the spherical surface [55- 57]. For each viewpoint, an image is taken by virtually simulating the shading process using OpenGL functions. Then some classification techniques [58] can be applied based on extracted features that discriminates image content from different viewpoints. To extract meaningful line drawings from a shaded image, the CLD method can be applied [51] An efficient line drawing clustering method by mimicking keyframe selection from a video clip is proposed in [57]. After clustering, each 3D object is represented by several representative line drawings. The number of representative line drawings adaptively depends on the model's complexity, e.g., for a sphere model, one representative circle is enough; for a complex model such as a camel, three or more representative line drawings are generated (Fig5b).

3.2 Memory coding as local descriptors

Image contents are usually characterized by local image descriptors. Several efficient local descriptors for line drawings have been proposed, such as the angular partitioning descriptor [59], the Gabor local line-based descriptor [60] and the circular histogram [57].

By considering that a line drawing is a hand-drawn rough black and white, an angular-spatial distribution of pixels is used in [59] to build a local descriptor that is rotation and scale invariant and robust against translation. Observing that a line drawing mainly consists of elongated lines on a constant background, Eitz et al. [60] proposes to use the Curvelet basis as the maximally sparse representation for such data. Liu et al. [57] observed that if a sketch or a line drawing depicts a person's mind, it is usually inaccurate and repeated sketching the same shape frequently shows local distortion like angular squeezing or stretching. Accordingly, they propose a circular histogram that only keeps the radial information in a local region of line drawings and is thus insensitive to angular variations.

The circular histogram proposed in [57] is computed as follows. First the minimum-area enclosing box B of all black pixels in a line drawing is computed (Fig.6a), which leads to a property of rotation invariance. Then a quasi-random point sequence is applied to sample 500 points inside B (Fig.6b). Centered at each sample point, a circular region C is located. The radius of the circle is one fifth of the diagonal length of B. The circular region C is further partitioned into 20 circular rings by partitioning r into 20 equal intervals and determining 20 concentric circles (Fig.6c). To build a circular histogram, each circular ring is served as a bin and its number is the number of black pixels fallen into that circular ring (Fig.6d). At this end, a line drawing is represented in a distributed way [4] and stored as 500 local circular histograms.

The local descriptors such as circular histograms can be regarded as a point in a high-dimensional feature space. For example, a circular histogram is a point in 20-dimensional space. Given *m* local descriptors sampled from *n* objects, some clustering techniques can be applied in the feature space such that these *m* points are partitioned into a small number *k* of clusters and each cluster is represented by a representative point. Then these *k* representative points are stored in memory as abstract local forms extracted from *n* objects.

Recently some works in neuroscience [61, 62] show that in addition to local descriptors, the shape skeleton information also play a role in memory encoding. In particular, the brain seems to identify an object by its parts and early computation of individual parts' skeletons leads to encoding of aggregation in a hierarchy of skeletons at later stages in the human vision system [62]. Toward this direction, an interesting line drawing generation method using skeleton information is presented in [63].

3.3 Learning in feature space

The abstract local forms should not be static in memory since human has a remarkable capacity of learning and the abstract local forms are evolved when more and more objects are seen. A continuous learning process is modeled in [4]. This learning model consists of two parts:

(1) The number *k* of abstract local forms in memory is evolved. When human sees more objects, *k* is increased with large possibility. When human did not see some objects for a long time, *k* is decreased with large possibility. This factor is modeled by a Poisson process.

(2) The abstract local forms themselves are evolved, e.g., when some objects are gradually forgotten and some novel objects are seen later. This factor is modeled by a continuous-time Markov chain.

These two-factor learning model was applied in [57]. A user study in line-drawing-based conceptual design showed that it models user's adaptive learning capacity well and can be used in a human-centered design environment.

The learning model in [57] needs hundreds of training data to obtain good performance. Another learning technique [64], called one-shot learning, can learn from only one or a very few examples of line drawings. This technique makes use of a global structural decomposition that decomposes a line drawing into a hierarchy of tokens, strokes and substrokes. This hierarchical structure can be described by the Backus normal form [65]. The learning and inference in [64] is modeled by a hierarchical Bayesian program.

3.4 Judgment using local descriptors

Most judgment tasks in computer vision are related to object recognition. A hypothesis in human vision is *recognition by component*

(RBC) [17]. The fundamental RBC hypothesis is that a small set of local components are generally invariant over viewing positions and are sufficient for diverse object representations and recognition. That is, any object is represented by a unique relation of a small set of local components. The abstract local forms stored in memory are consistent with the RBC hypothesis.

To use these evolving abstract local forms for efficient object recognition, the classic bag-of-feature or bag-of-word approach [66] in image retrieval can be applied [60, 57]. The bag-of-feature approach makes use of local image descriptors such that partial matching can be efficiently done, which is also reliable to global deformation inherent in different line drawings.

In terms of k abstract local forms $\{f_1, f_2, \ldots, f_k\}$, each form f_i functions like a visual word in the bag-of-feature approach. Given any line drawing LD with 500 local circular histograms, it is encoded by LD = $(n_1 f_1, n_2 f_2, \ldots, n_k f_k)$, where n_i is the number of circular histograms (i.e., points in a feature space) fallen into the cluster of f_i. Sometimes, inverse document frequency [67] can be applied to adjust the weight n_i. Given the vector encoding LD(I) of an line drawing I query and the vector encoding LD(L) of any line drawing L in a database, the similarity between I and L can be obtained by a metric from an inner product $\langle LD(I), LD(L) \rangle$. The object (represented by line drawings of representative views) with the highest similarity value is recognized for the line drawing queries.

3.5 3D reconstruction from line drawings

In addition to the computational cognitive process of line drawings as presented in Sections 3.1 to 3.4, there is another important stream in computer vision research, called machine interpretation of line drawings [68]. Its main target is to investigate a computational mechanism for reconstructing 3D structures from 2D line drawings. Basically the 3D structure under investigation is in the form of polyhedrons. We emphasize that this is a very difficult ill-posed problems, since (1) solid polyhedrons bounded by planar faces cannot be fully specified by line drawings, since they are only single-view pictures of 2D projection, and (2) human perception seems have infinite understandings of line drawings and thus some priori knowledge as well as specific domain knowledge must be used. We also note that this kind of computational intelligence for 3D reconstruction may be quite different from computational cognitive process employed in human perception. Some excellent surveys on early work of 3D reconstruction from line drawings can be found in [68, 69].

4. Intelligent processing and applications of line drawings

Since line drawings can naturally represent the human mind, they are widely used in diverse application of visual media. In this section, we summary some state-of-the-art applications for processing visual media including images, videos and 3D digital geometry. Equipped with sketching interfaces in touch devices such as Tablets or iPads, these line-drawing-based applications show some intelligence since the users' indent can be addressed naturally and fluently by mimicking traditional paper-and-pencil-based 2D sketches.

4.1 Interaction with image data

It would be fantastic if we can convert our thought into photorealistic images using some computer program. Sketch2Photo [70] is such a technique. It novelly combines two methods: one is line drawing that easily captures human thought but with limited realism, the other is photomontage that uses massive existing photographs to compose a novel image.

Sketch2Photo works in the following way. First, to better capture a user's mind, a user draws a free-hand line drawing together with some necessary text labels. Second, Each scene item as well as the background in the line drawing was searched using the text label in the internet. Third, to exclude undesirable images, the search results are filtered by segmenting searched images into scene items and matching the shape of scene items with line drawings. Fourth, to seamlessly compose the filtered images, an image blending technique is used. Finally, several candidates of image compositions are automatically generated and recommended to the user. Three examples in Sketch2Photo are shown in Fig.7.

Sketch-based image search plays an important role in a large variety of image applications including Sketch2Photo [70]. Although it had been extensively studied two decades ago, the research of sketch-based image search is undergoing a renaissance due to the explosion of web images and the popularity of touch devices. A novel index structure with a contour-based matching algorithm was proposed in [71] to assess the similarity between a sketch query and natural images in a database. This method can indexes 2.1 million Flickr images with 6.5GB memory (suitable to store in a common server) and supports real-time response by returning search results around 1 second.

4.2 Interaction with video data

Nowadays, with the explosion of video data on the internet, a concise form that can assist human to quick grasp the gist in a long video is much desired. Such a concise form is usually referred to video summarization or video abstraction in computer vision and multimedia societies [72]. Previously keyframes and keywords are widely used as efficient video abstraction forms. In recent work [42, 73], a user study showed that line drawings generally outperform keyframes and keywords as an abstraction form in video annotations for human understanding. See Fig.8 for an example. According to this result, an efficient video authoring tool, which creates from a collection of video clips a context-aware, interactive video representation, was presented in [73]. This tool utilizes a sketch-based two-layer representation (called SSG). One layer in SSG uses line drawings to visualize scene information. The other layer uses a graph to represent and edit the narrative structure of the authoring interactive video representation. Fig.9a shows an example. To reuse the knowledge in existing SSGs, a new SSG can be created by matching the line drawings and combining the graphs in existing SSGs (Fig.9b)

In a further work [42], a single-layer sketch graph was proposed for efficient organization of large-scale video clips. Line drawings are also used as an efficient way to represent the user's mind for organization purpose and two kinds of knowledge are considered. Each node in a sketch graph is a line drawing that reflects declarative knowledge, that is, some factual information. Each edge in a sketch graph is a free-hand stroke that reflects procedural knowledge, that is, skills in performing some tasks that is usually dynamic. User studies in [42] showed that sketch graphs can better represent the organization structure in a large-scale video clips than a simple combination of keyframes and keywords (Fig.10), and are more efficient than SSG in [73] in terms of scalability.

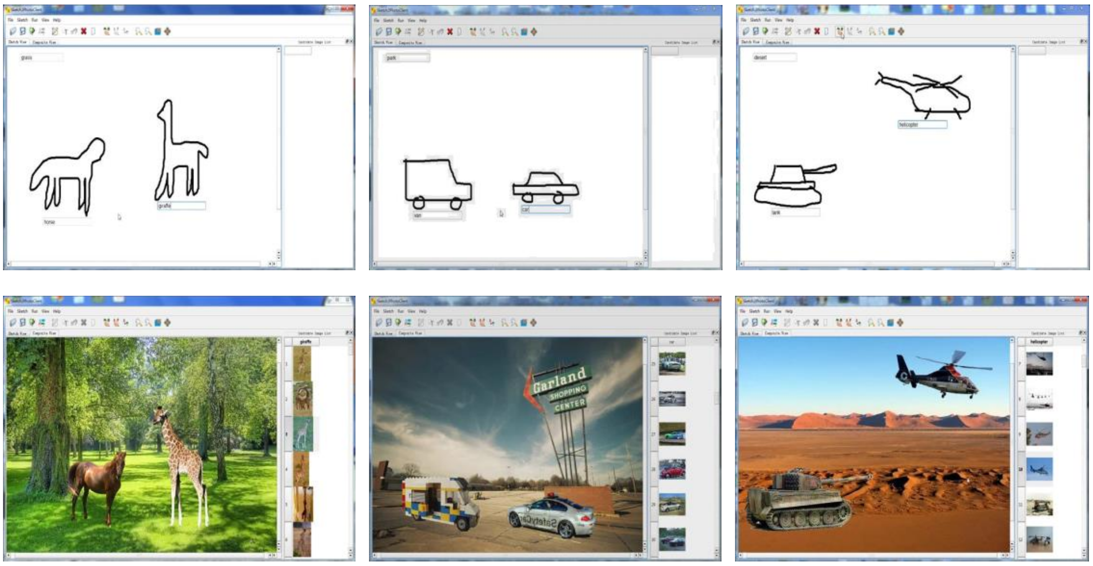


Fig.7 Three examples of Sketch2Photo [70]. Top row is input and bottom row is output.

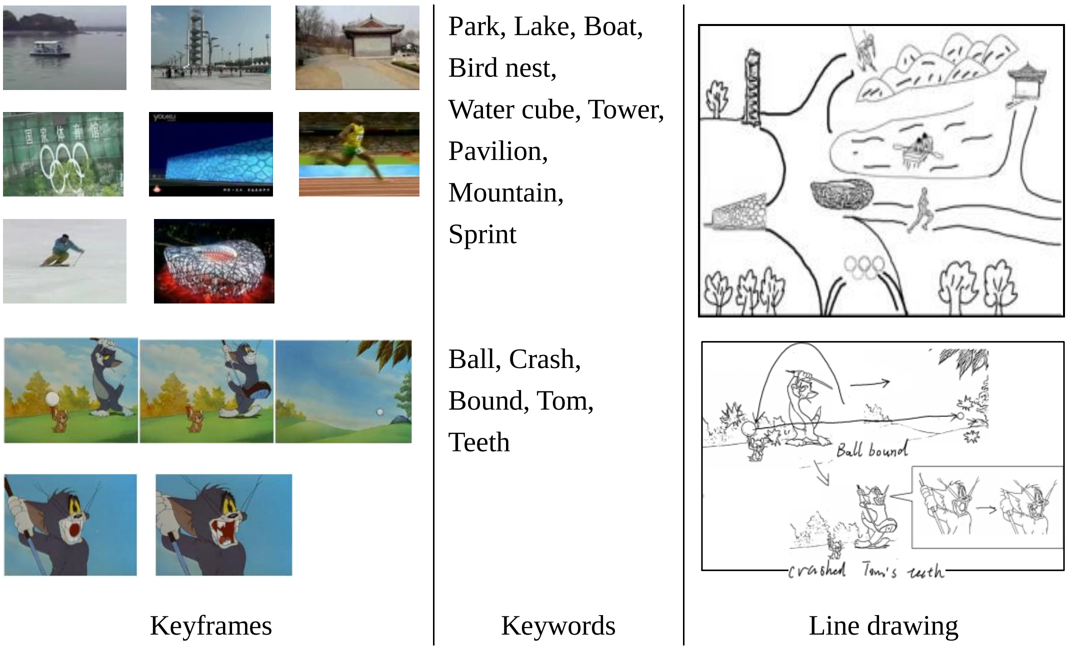
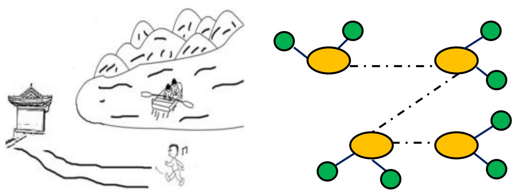
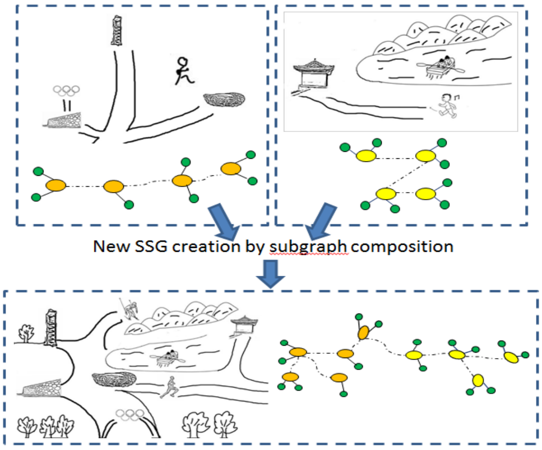


Fig.8 Three abstraction forms (keyframes, keywords and line drawings) for video clips

Both multimedia authoring [73] and video organization [42] rely on an efficient sketch-based video retrieval. Since line drawings or sketches are good abstraction forms of videos, they had been considered as an effective media for video retrieval. A content based video retrieval system (CBVR), driven by free-hand sketch queries, was proposed in [74]. This system encodes the shape and motion of a sketched object. Considering the sketches drawn for CBVR are often imprecision with respect to both appearance and motion, an autoregressive model was proposed for video clips. The sketch queries are then matched to clips for objects and their movement. In [73], a bag-of-visual-words method is applied for sketch-based video retrieval. This method decomposes video clips into shots and keyframes, and depicts sketch queries using the circular histograms as shown in Fig.6. By converting each keyframe into a sketch using



(a) Two layers in a sketch-based two-layer representation (SSG)



(b) New SSG creation by reusing existing SSGs

Fig.9 The sketch-based two-layer representation (SSG) for video clips in [73]

the CLD method, the sketch-based video retrieval is transformed to a sketch-to-sketch retrieval problem. Both sketch-based video retrieval methods [73] [74] cannot be scalable to a large-scale video database. In [42], an index structure is proposed with the circular histogram descriptor, which leads to an efficient large-scale sketch-based search in a database of more than ten thousands of video clips.

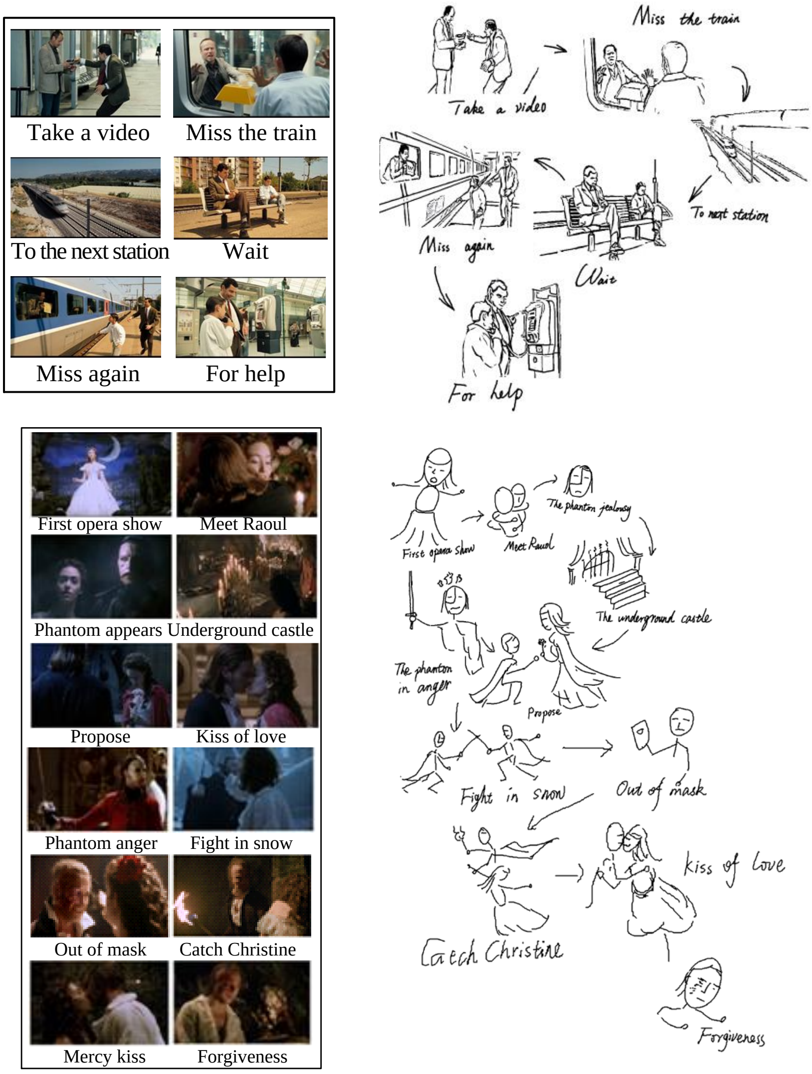
4.3 Interaction with 3D digital geometry data

3D free-form geometric models are usually difficult to create by interaction with computers. Traditional powerful commercial software such as 3DMax and Maya all uses the WIMP (windows, icons, menus, pointer) interface. This interface is well known far from natural and efficient in terms of interaction performance. Sketching interface had been studied for creating 3D geometry by sketching some simple line drawings [75-77]. Teddy [75] is a classic system that uses sketching interface to design 3D free-form shape from free-form strokes. Sketching with gesture operations can naturally express the user's design intent. However, the number of gestures that are used for free-form shape modeling cannot be large due to limited capacity of human working memory. Error-prone behavior also restricts the applications of sketching based 3D shape modeling. By extending freeform stokes to 3D control curves, a progress is made in a FiberMesh system [76] that can create much more sophisticated models. A new sketching form, called *editable sketching curves*, was further proposed in [77] that combine the advantages of natural expression of free-form strokes and the controllability of B-spline curves. To speed up sketch-based operations for a real-time performance, GPU acceleration technique has been studied in [78].

An interactive toy design system called EasyToy was built up with editable sketching curves [77]. In EasyToy, both shape and color are manipulated by simple sketching tools, that is, five tools for shape (extrusion, Boolean, disk deformation, skeleton deformation and bounding box deformation; See Fig.11) and four tools for color (pick a color, paint with sketching curve, fill in a local region and paint with strokes; See Fig.12). EasyToy can construct sophisticated models whose complexity is comparable with those by professional systems such as 3DMax and Maya (Fig.13).

5. Future directions

Despite significant progress of research on line drawings over the last two decades, several potential avenues still exist for future

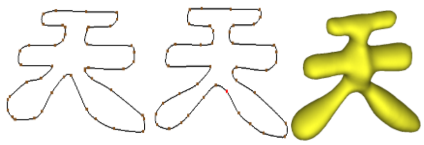


Keyframe+keywords Sketch graph

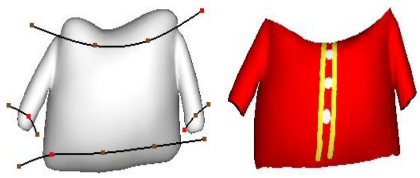
Fig.10 Organization of large-scale video clips using keyframes+keywords, or using sketch graph

studies. Below we summarize five research directions.

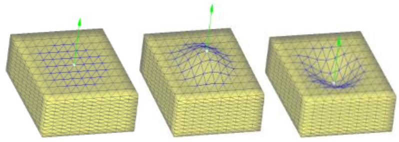
1. *Cognition mechanism*. Although fMRI and ERP techniques had been applied to cognition mechanism research by comparing natural photographs and line drawings, we are still far from clear what are the underlying neural codes of human vision. For example, does the spectral magnitude or phase play a role in line drawing understanding? The answer to this question relates to global (long strokes) or local (short strokes) structures in a line drawing. We also do not know why artists understand (possibly implicit) the neuro codes of vision by having the capacity to choose the right lines for depicting depth, occlusion, shadow or motion (Figs. 1, 2c and 4). More advanced brain studies should be conducted to answer these questions.
2. *A computer program to extract aesthetic line drawing from images*. So far only artists can select right lines in images for generating a meaningful and aesthetic line drawing. Even state-of-the-art computer algorithms can only be applied in restricted cases. For example, for the



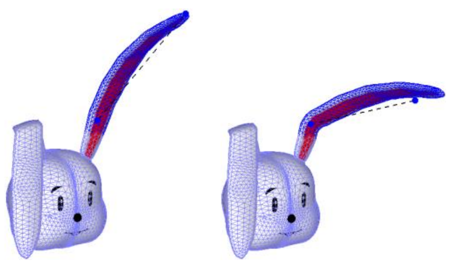
Extrusion with editable sketching curves



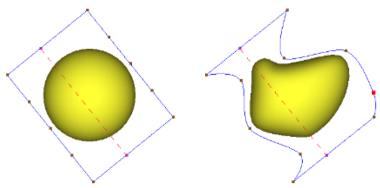
Boolean operations (cut and paste)



Disk deformation



Skeleton deformation



Bounding box deformation

Fig.11 Five sketching operators for shape manipulation in EasyToy [77]

classic CLD algorithm [51], we already know that 1) it works well for cartoon pictures, and 2) it works very badly on low resolution images such as the keyframes extracted from old movies, 3) it works unstably for other natural images (Fig. 14). Then developing a good computer algorithm that can mimic artists' performance will be definitely

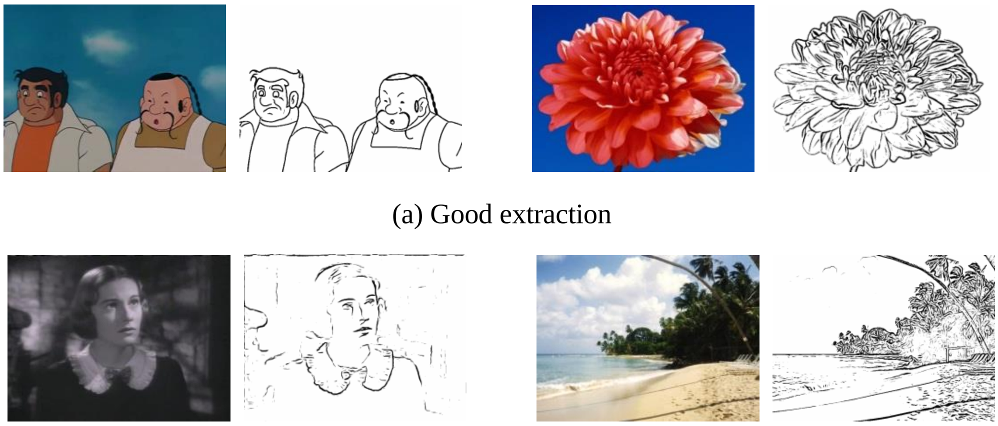
Fig.12 Four sketching operators for color manipulation in EasyToy [77]. (a) Pick a color. (b) Paint with an editable sketching curve. (c) Fill in a local region with a color. (d) Paint with strokes.

useful in many applications such as video abstraction, organization and multimedia authoring. Many image techniques such as 2.1D sketch generation [79], salient region detection, foreground/background separation and feature enhancement, may be needed to be considered in a combinatorial framework.

3. *Better learning methodology*. Human can learn and recognize hand written characters and line drawings quickly and accurately. The one-shot learning technique [64] made a step in modeling such a learning capacity. However, the complexity of the shape hierarchy in [64] is low and if a complex line drawing is given, the computational complexity increases explosively in the hierarchical Bayesian model. Developing an efficient learning methodology for complex line drawings is still under exploration.

4. *Scalability to large databases*. Line drawing or sketch based visual media retrieval has attracted considerable attention and many good algorithms had been proposed. In these algorithms, however, few of them have good scalability to fit in a large-scale database. Some exceptional work includes image

Fig.13 Some models (bottom row) created from line drawings (top row) in EasyToy [77]



(a) Good extraction

(b) Bad extraction

Fig.14 The performance of CLD algorithm [51] on different types of images.

Fig.15 Different sketching interface and devices with touchable screen for line drawing applications

search [71], video search [42] and 3D geometry [80] in a large scale. The explosive growth of internet visual media has motivated the great demand for more search/retrieval techniques with good scalability for line drawing queries.

5. *More exciting intelligent applications in visual media*. Line drawings are natural to represent human mind, which are also naturally incorporated in devices with touchable screens. With the advance of hardware development for touch devices

(Fig.15), more exciting intelligent applications are expected.

6. Conclusions

In this paper we present a comprehensive survey on cognitive mechanism of line drawings and its applications in intelligent process of visual media. Several key issues in cognitive process of line drawings in human brain are discussed, including surface versus edge-based representation of visual perception, features used in the perception of line drawings, human brain activity and time course of visual perception, and semantic access to conceptual representation. Based on these neurophysiological findings, intelligent processing of line drawings on visual media, including image, video and 3D digital geometry, is discussed with various applications.

Line drawing or sketch based visual media processing, as well as its related cognitive study, is an active and multidisciplinary area of research. This survey is not meant to complete, especially for a large variety of applications in the intelligent processing of visual media. Nevertheless, we hope this survey can provide some information for researchers to review the past developments and identify possible directions for future research on cognitive mechanism and intelligent processing of line drawings.

Acknowledgement

This work was supported by the Natural Science Foundation of China (61322206, 61272228), the National Basic Research Program of China (2011CB302200), Beijing Higher Institution Engineering Research Center of Visual Media Intelligent Processing and Security, Tsinghua National Lab for Information Science and Technology Cross-Discipline Foundation, and Tsinghua University Initiative Scientific Research Program (20131089252).

Reference

1. Fischetti M. Computers vs. Brains. Scientific American, 2011, 305(5): 104-104.
2. Fu X L, Cai L H, Liu Y, Jia J, Chen W F, Zhang Y, Zhao G Z, Liu Y J, Wu C X. A computational cognition model of perception, memory and judgment. Science in China (Series F: Information Sciences), Vol. 57, No. 3, Article No. 032114(1-15), 2014.
3. Ullman S. High-level vision: Object recognition and visual cognition. MIT press, 2000.
4. Liu Y J, Fu Q F, Liu Y, Fu X L. A distributed computational cognitive model for object recognition. Science China Information Sciences, 2013, 56(9): 1-13.
5. Riddoch M, Humphreys G. Object Recognition. In B. Rapp (Ed.), Handbook of Cognitive Neuropsychology. Hove: Psychology Press.
6. Canny J. A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1986 (6): 679-698.
7. Cole F, Golovinskiy A, Limpaecher A, et al. Where do people draw lines? ACM Transactions on Graphics (TOG). ACM, 2008, 27(3): 88..
8. Sayim B, Cavanagh P. What Line Drawings Reveal About the Visual Brain. Frontiers in Human Neuroscience, Volume 5, Article No. 118, 2011.
9. Yonas A, Arterberry ME. Infants perceive spatial structure specified by line junctions. Perception. 1994;23(12):1427-35.
10. Itakura S. Recognition of line-drawing representations by a chimpanzee (Pan troglodytes). The Journal of general psychology, 1994, 121(3): 189-197.

11. Cole F, Sanik K, DeCarlo D, et al. How well do line drawings depict shape? ACM Transactions on Graphics (TOG). ACM, 2009, 28(3): 28.
12. Koenderink J J, Van Doorn A J, Kappers A M L. Surface perception in pictures. Perception & Psychophysics, 1992, 52(5): 487-496.
13. Koenderink J J, van Doorn A J, Kappers A M L, et al. Ambiguity and themental eye'in pictorial relief. PERCEPTION-LONDON-, 2001, 30(4): 431-448.
14. David Marr. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. The MIT Press, July 9, 2010.
15. Watt R J, Rogers B J. Human vision and cognitive science. Research Directions in Cognitive Science: A European Perspective, 1989, 1: 9-22.
16. Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision. Cengage Learning, 2014.
17. Biederman I. Recognition-by-components: a theory of human image understanding. Psychological review, 1987, 94(2): 115.
18. Biederman I, Ju G. Surface versus edge-based determinants of visual recognition. Cognitive psychology, 1988, 20(1): 38-64.
19. Fu Q, Liu Y J, Chen W, et al. The time course of natural scene categorization in human brain: simple line-drawings vs. color photographs. Journal of Vision, 2013, 13(9): 1060-1060.
20. Del Viva M M, Punzi G, Benedetti D. Information and Perception of Meaningful Patterns. PLoS One, 2013, 8(7): e69154.
21. Morgan M J. Features and the 'primal sketch'. Vision research, 2011, 51(7): 738-753.
22. Delorme A, Richard G, Fabre-Thorpe M. Key visual features for rapid categorization of animals in natural scenes. Frontiers in psychology, 2010, 1: 21.
23. Naber M, Hilger M, Einhäuser W. Animal detection and identification in natural scenes: image statistics and emotional valence. Journal of vision, 2012, 12(1): 25.
24. Derrington A M, Krauskopf J, Lennie P. Chromatic mechanisms in lateral geniculate nucleus of macaque. The Journal of Physiology, 1984, 357(1): 241-265.
25. Campbell F W, Robson J G. Application of Fourier analysis to the visibility of gratings. The Journal of physiology, 1968, 197(3): 551.
26. Joubert O R, Rousselet G A, Fabre-Thorpe M, et al. Rapid visual categorization of natural scene contexts with equalized amplitude spectrum and increasing phase noise. Journal of Vision, 2009, 9(1): 2.
27. Jolicoeur P. The time to name disoriented natural objects. Memory & Cognition, 1985, 13(4): 289-303.
28. Farah M J, Hammond K M. Mental rotation and orientation-invariant object recognition: Dissociable processes. Cognition, 1988, 29(1): 29-46.
29. Westheimer G. The Fourier theory of vision. PERCEPTION-LONDON-, 2001, 30(5): 531-542.
30. Oppenheim A V, Lim J S. The importance of phase in signals. In: Proceedings of the IEEE, 1981, 69(5): 529-541.
31. Piotrowski L N, Campbell F W. A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase[J]. Perception, 1982.
32. Guyader N, Chauvin A, Peyrin C, et al. Image phase or amplitude? Rapid scene categorization is an amplitude-based process. Comptes Rendus Biologies, 2004, 327(4): 313-318.
33. Chen W F, Liang J, Liu Y J, Fu Q F, Fu X L. Rapid natural scene categorization of line drawings is less influenced by amplitude spectra: Evidence from a subliminal perception study. ASSC 18: Poster Session (Association for the Scientific Study of

Consciousness).
34. Morrone M C, Burr D C. Feature detection in human vision: A phase-dependent energy model. In: Proceedings of the Royal Society of London. Series B, biological sciences, 1988: 221-245.
35. Devlin H, Tracey I, Johansen-Berg H, Clare S. What is Functional Magnetic Resonance Imaging (fMRI)? FMRIB Centre, Department of Clinical Neurology, University of Oxford.
36. Walther D B, Caddigan E, L F F, et al. Natural scene categories revealed in distributed patterns of activity in the human brain. The Journal of Neuroscience, 2009, 29(34): 10573-10581.
37. Walther D B, Chai B, Caddigan E, et al. Simple line drawings suffice for functional MRI decoding of natural scene categories. In: Proceedings of the National Academy of Sciences, 2011, 108(23): 9661-9666.
38. Kim S G, Richter W, Uğurbil K. Limitations of temporal resolution in functional MRI. Magnetic Resonance in Medicine, 1997, 37(4): 631-636.
39. Vanrullen R, Thorpe S J. The time course of visual processing: from early perception to decision-making. Journal of cognitive neuroscience, 2001, 13(4): 454-461.
40. Johnson J S, Olshausen B A. Timecourse of neural signatures of object recognition. Journal of Vision, 2003, 3(7): 4.
41. Fu Q F, Liu Y J, Dienes Z, Wu J H, Chen W F, Fu X L. Different time courses of natural scene categorization of color photographs and line-drawings: Evidence from event-related potentials. Submitted for publication, 2014.
42. Liu Y J, Ma C X, Fu Q, et al. A Sketch-Based Approach for Interactive Organization of Video Clips. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2014, 11(1): 2.
43. Snodgrass J G, Vanderwart M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. Journal of experimental psychology: Human learning and memory, 1980, 6(2): 174.
44. Magnié M N, Besson M, Poncet M, et al. The Snodgrass and Vanderwart set revisited: Norms for object manipulability and for pictorial ambiguity of objects, chimeric objects, and nonobjects. Journal of Clinical and Experimental Neuropsychology, 2003, 25(4): 521-560.
45. Rossion B, Pourtois G. Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. PERCEPTION-LONDON-, 2004, 33(2): 217-236.
46. Viggiano M P, Vannucci M, Righi S. A new standardized set of ecological pictures for experimental and clinical research on visual object processing. Cortex, 2004, 40(3): 491-509.
47. Abel S, Weiller C, Huber W, et al. Neural underpinnings for model-oriented therapy of aphasic word production. Neuropsychologia, 2014, 57: 154-165.
48. Janssen N, Carreiras M, Barber H A. Electrophysiological effects of semantic context in picture and word naming. Neuroimage, 2011, 57(3): 1243-1250.
49. Schnur T T. The persistence of cumulative semantic interference during naming. Journal of Memory and Language, 2014, 75: 27-44.
50. Strijkers K, Holcomb P J, Costa A. Conscious intention to speak proactively facilitates lexical access during overt object naming. Journal of memory and language, 2011, 65(4): 345-362.
51. Kang H, Lee S, Chui C. Coherent line drawing. In: Proceedings of 5th International Symposium on Non-photorealistic Animation and Rendering. New York: ACM, 2007. 43–50
52. Arbel T, Ferrie F P. Viewpoint selection by navigation through entropy maps.Computer

Vision, 1999. In: Proceedings of the Seventh IEEE International Conference on. IEEE, 1999, 1: 248-254.
53. Laporte C, Arbel T. Efficient discriminant viewpoint selection for active bayesian recognition. International Journal of Computer Vision, 2006, 68(3): 267-287.
54. Secord A, Lu J, Finkelstein A, et al. Perceptual models of viewpoint preference. ACM Transactions on Graphics (TOG), 2011, 30(5): 109.
55. Chen D Y, Tian X P, Shen Y T, et al. On visual similarity based 3D model retrieval.Computer graphics forum. Blackwell Publishing, Inc, 2003, 22(3): 223-232.
56. Cyr C M, Kimia B B. 3D object recognition using shape similiarity-based aspect graph.Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. IEEE, 2001, 1: 254-261.
57. Liu Y J, Luo X, Joneja A, Ma C X, Fu X L, Song D W (2013) User-adaptive sketch-based 3-D CAD model retrieval. IEEE Transactions on Automation Science and Engineering, Vol. 10, No. 3, pp. 783-795, 2013.
58. Duda R O, Hart P E, Stork D G. Pattern classification. John Wiley & Sons,, 1999.
59. Chalechale A, Naghdy G, Mertins A. Sketch-based image matching using angular partitioning. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2005, 35(1): 28-41.
60. Eitz M, Richter R, Boubekeur T, et al. Sketch-based shape retrieval. ACM Trans. Graph., 2012, 31(4): 31.
61. Hung C C, Carlson E T, Connor C E. Medial axis shape coding in macaque inferotemporal cortex. Neuron, 2012, 74(6): 1099-1113.
62. Lescroart M D, Biederman I. Cortical representation of medial axis structure. Cerebral cortex, 2013, 23(3): 629-637.
63. Li Z, Qin S, Jin X. Skeleton-enhanced line drawings for 3D models[J]. Graphical Models, 2014, 76(6): 620-632.
61. Lake B M, Salakhutdinov R, Tenenbaum J. One-shot learning by inverting a compositional causal process.Advances in neural information processing systems. 2013: 2526-2534.
62. Ma C X, Liu Y J, Yang H Y, Teng D X, Wang H A, Dai G Z (2011) KnitSketch: a sketch pad for conceptual design of 2D garment patterns. IEEE Transactions on Automation Science and Engineering, Vol. 8, No. 2, pp. 431-437, 2011.
66. Sivic J, Zisserman A. Video Google: A text retrieval approach to object matching in videos.Computer Vision, 2003. In:Proceedings. Ninth IEEE International Conference on. IEEE, 2003: 1470-1477.
67. Baeza-Yates R, Ribeiro-Neto B. Modern information retrieval. New York: ACM press, 1999.
68. Sugihara K. Machine interpretation of line drawings. Cambridge: MIT press, 1986.
69. Hoffman D D. Visual intelligence: How we create what we see. WW Norton & Company, 2000.
70. Chen T, Cheng M M, Tan P, et al. Sketch2Photo: internet image montage.ACM Transactions on Graphics (TOG). ACM, 2009, 28(5): 124.
71. Cao Y, Wang C H, Zhang L Q, Zhang L. Edgel index for large-scale sketch-based image search. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 761-768, 2011.
72. Truong B T, Venkatesh S. Video abstraction: A systematic review and classification. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), 2007, 3(1): 3.
73. Ma C X, Liu Y J, Wang H A, Teng D X, Dai G Z.(2012) Sketch-based annotation and visualization in video authoring. IEEE Transactions on Multimedia, Vol. 14, No. 4, pp.1153-1165, 2012.

74. Collomosse J P, McNeill G, Qian Y. Storyboard sketches for content based video retrieval. Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009: 245-252.
75. Igarashi T, Matsuoka S, Tanaka H. Teddy: a sketching interface for 3D freeform design. ACM Siggraph 2007 courses. ACM, 2007: 21.
76. Nealen A, Igarashi T, Sorkine O, et al. FiberMesh: designing freeform surfaces with 3D curves.ACM Transactions on Graphics (TOG). ACM, 2007, 26(3): 41.
77. Liu Y J, Ma C X, Zhang D L. EasyToy: plush toy design using editable sketching curves. IEEE Computer Graphics and Applications, 2011, 31(2): 49-57.
78. Zhu X, Jin X, Liu S, et al. Analytical solutions for sketch-based convolution surface modeling on the gpu. The Visual Computer, 2012, 28(11): 1115-1125. 5
79. Yu C C, Liu Y J, Wu T F, Li K Y, Fu X L (2014) A global energy optimization framework for 2.1D sketch extraction from monocular images. Graphical Models, Vol. 76, No. 5, pp. 507-521, 2014.
80. Liu Y J, Zhang J B, Hou J C, et al. Cylinder Detection in Large-Scale Point Cloud of Pipeline Plant. Visualization and Computer Graphics, IEEE Transactions on, 2013, 19(10): 1700-1707.

Yong-Jin Liu received his B.Eng. degree from Tianjin University, China, in 1998, and his Ph.D degree from the Hong Kong University of Science and Technology, Hong Kong, in 2004. He is now an Associate Professor with the Tsinghua National Laboratory for Information

Science and Technology, Department of Computer Science and Technology, Tsinghua University, China. His research interests include pattern recognition, computer graphics, computational geometry and computer-aided design. He is a member of IEEE, a member of IEEE Computer Society and IEEE Communications Society. More research topics about him can be found at http://cg.cs.tsinghua.edu.cn/people/~Yongjin/yongjin.htm

Minjing YU received her B.Eng. degree from Wuhan University, China, and she is now a PhD student at Department of Computer Science and Technology, Tsinghua University, China. Her research interests include image processing, computer graphics and cognitive science.

Qiufang Fu received her PhD degree from Institute of Psychology, Chinese Academy of Sciences, in 2006. She is now an associate professor in Institute of Psychology, CAS, with interests in implicit learning and unconscious knowledge. She intends to explore the neural and cognitive mechanisms responsible for the dissociation of implicit and explicit processes.

Wenfeng Chen received his B.S. degree from Peking University, China, in 1999, and his PhD degree from Institute of Psychology, Chinese Academy of Sciences, China in 2005. He is now an associate professor of psychology in the Institute of Psychology, Chinese Academy of Sciences, and a guest Associate Editor of Frontiers in Psychology. He is interested in face learning and recognition, visual attention, emotion and cognitive modelling.

Ye Liu received her PhD from Institute of Psychology, Chinese Academy of Sciences, China in 2005. She is now an associate professor of psychology in the Institute of Psychology, Chinese Academy of Sciences. She is interested in semantic knowledge representation and semantic processing, metaphor comprehension, and affective computing.

Lexing Xie is Senior Lecturer in the Research School of Computer Science at the Australian National University. She was research staff member at IBM T.J. Watson Research Center in New York from 2005 to 2010, and adjunct assistant professor at Columbia University 2007-2009. She received B.S. from Tsinghua University, Beijing, China, and M.S. and Ph.D. degrees from Columbia University, all in Electrical Engineering. Her research interests are in applied machine learning, multimedia, social media. Her research has received five best student paper and best paper awards between 2002 and 2011. She currently serves an associate editor of IEEE Transactions on Multimedia, and ACM Transactions on Multimedia Computing, Communications and Applications.