

시각 예제에 의한 질의: 시각정보 검색지원을 위한 이미지 질의 패러다임의 유용성 비교 연구

Query by Visual Example: A Comparative Study of the Efficacy of Image Query Paradigms in Supporting Visual Information Retrieval

콜린 벤터스*
Colin C. Venters

차 례

1. Introduction	5. Discussion
2. Research Background	6. Conclusion
3. Experiment	· References
4. Results	

초 록

시각적 실례에 의한 질의는 내용기반 이미지 검색 환경에서 질의 표현을 위한 중요한 질의 패러다임이다. 이 이미지 및 스케치에 의한 질의는 질의표현을 가능하게 하는 방법으로서 오랫동안 알려졌다. 하지만 이 방법이 질의의 쉽게 작성하는 데 얼마나 도움을 주는지에 대한 효율성에 대한 실험적 입증은 아직 미미하다. 정보검색시스템에 표현하는 탐색자의 능력은 검색과정의 기본이다. 이 연구의 목적은 탐색자의 정보 문제와 효율적이고도 효과적인 시각적 질의 작성을 지원하기 위해 필요한 질의 방법들 간의 지식 격차의 원인이 되는 다양한 정보 요구를 지원하는 데 있어서 유용성 실험을 통해 이미지에 의한 질의와 스케치방법에 의한 질의 조사하기 위한 이었다. 본 연구 결과는 이미지에 의한 질의가 시각적 질의 작성에 실행 가능한 접근방식임을 제시한다. 반면에, 본 연구결과를 통해 탐색자의 정보 문제와 시각적인 질의 작성에 도움을 주는 스케치 패러다임에 의한 질의표현 능력 간에 상당한 불일치가 있다는 것을 알 수 있다. 효율(시간)과 유효성(오류)에 초점을 둔 유용성 실험결과와 이용자의 만족도는 큰 차이점이 있다고 보여준다($p<0.001$). 이는 다음 세 가지 측정(시간, 오류, 이용자의 만족도)에 대한 두 가지 질의 방식(이미지에 의한 질의, 스케치에 의한 질의) 사이에서 나타난 시간($Z=-3.597$, $p<0.001$), 오류($Z=-3.317$, $p<0.001$), 그리고 만족도($Z=-10.223$, $p<0.001$)에서 드러난다. 본 연구결과는 또한 질의도구를 참가자가 인지하는 유용성에 큰 차이가 있다는 것을 보여준다($Z=-4.672$, $p<0.001$).

키 워 드

내용기반 이미지 검색, 이미지 검색, 이미지에 의한 질의, 시각사례에 의한 질의, 스케치에 의한 질의, 시각 정보검색, 유용성

* 경북대학교 문헌정보학과 조교수
(Assistant Professor, Dept. of Library and Information Science, Kyungpook National University, venters@knu.ac.kr)
 • 논문접수일자: 2011년 5월 24일
 • 최종심사(수정)일자: 2011년 5월 30일
 • 게재확정일자: 2011년 6월 1일

ABSTRACT

Query by visual example is the principal query paradigm for expressing queries in a content-based image retrieval environment. Query by image and query by sketch have long been purported as being viable methods of query formulation yet there is little empirical evidence to support their efficacy in facilitating query formulation. The ability of the searcher to express their information problem to an information retrieval system is fundamental to the retrieval process. The aim of this research was to investigate the query by image and query by sketch methods in supporting a range of information problems through a usability experiment in order to contribute to the gap in knowledge regarding the relationship between searchers' information problems and the query methods required to support efficient and effective visual query formulation. The results of the experiment suggest that query by image is a viable approach to visual query formulation. In contrast, the results strongly suggest that there is a significant mismatch between the searchers information problems and the expressive power of the query by sketch paradigm in supporting visual query formulation. The results of a usability experiment focusing on efficiency (time), effectiveness (errors) and user satisfaction show that there was a significant difference, $p < 0.001$, between the two query methods on all three measures: time ($Z = -3.597$, $p < 0.001$), errors ($Z = -3.317$, $p < 0.001$), and satisfaction ($Z = -10.223$, $p < 0.001$). The results also show that there was a significant difference in participants perceived usefulness of the query tools $Z = -4.672$, $p < 0.001$.

KEYWORDS

Content-based Image Retrieval, Image Retrieval, Query by Image, Query by Visual Example, Query by Sketch, Visual Information Retrieval, Usability

1. Introduction

Advances in technology have been the catalyst for an unprecedented increase in the volume of institutional and personal image data. However, while there have been significant technological advances with image data capture and storage, developments in effective image retrieval have not kept pace. To address this, research into the retrieval of image data in the last two decades has focused

primarily on the use of visual properties and characteristics; commonly referred to as content-based image retrieval.

Within this field, a number of key research areas have been identified that are advocated as crucial to its advancement (Gudivada and Raghaven 1995; Rui, Huang, and Chang 1999; Hanjalic, Sebe, and Chang 2006; Datta et. al. 2008). Despite a long recognition as a key area of research, query interfaces to support effective information retrieval including their

expressive power and the impact of the query on retrieval performance, especially with real end users, has still to receive serious attention by the research community(Flickner et. al. 1997). Jaimes, Sebe and Gatica-Perez(2006) argue that in order to gain acceptance, emerging image retrieval systems need to acquire a human-centered perspective.

Jaimes and Chang(2002) state that content-based image retrieval systems should consist of three basic components at the highest level of abstraction: the database, the indexing information, and the user interface. When users' conduct searches, they interact with the system to express their information requirements through a user interface. They suggest that two important aspects of the user interface are its *expressive power* and its *ease of use*; where expressive power relates to what can be expressed using the language and ease of use is concerned with how difficult it is for the user to formulate the desired query using the language. This paper reports the results of an investigation that aimed to explore the *efficacy* of the *query by image* and *query by sketch* methods in supporting a range of information problems. Section two provides the background to the research outlining the query paradigms and related experimental evaluation studies. Section three

outlines the methods and procedures used in the experiment. Section four presents the results that emerged from the experiment. Section five discusses the implications of the results and draws conclusions.

2. Research Background

The origins of content-based image retrieval lie in the fields of artificial intelligence, computer vision, image processing, pattern recognition, and signal processing. Advancement of the approach has been attributed to the early experiments conducted by Kato (1991) into the automatic retrieval of images by colour and shape feature. The process involves a direct matching operation between a query image and a database of stored images. For each image a feature vector is computed for its unique visual properties

$$f(f_1^I, f_2^I, f_3^I, \dots, f_d^I)$$

〈Equation 1〉 Visual Feature Representation

where f is the feature vector, I is the representation of the image content, and d is the dimensionality. Similarity is calculated by comparing the feature vector of the query image against the feature vectors of images stored

in a database using a distance model such as *Euclidean distance*. Other commonly used metric distance functions include Kullback-Leibler divergence, Mahalanobis distance, and min-kowsky distance. The most common image matching feature characteristics are colour, shape, and texture. A number of review papers have surveyed the feature and matching techniques utilized in the process of content-based image retrieval and readers are directed to those publications for detailed descriptions of the techniques and approaches used (Idris and Panchanathan 1997; Del Bimbo 1999; Lew and Huang 2001; Castelli and Bergman 2002; Feng, Siu, and Zhang 2003). Nevertheless, the ability to provide effective and efficient image retrieval using this approach has proved challenging and its efficacy as a primary retrieval method is open to debate.

2.1 Query by Visual Example

Image retrieval is the process of identifying relevant images from a database that satisfies an information need. In the context of this research, an information need is a requirement to find relevant information in response to a searcher's specific information problem. This is formally expressed as a query Q_n using the input language provided by the retrieval

system. To provide both structure and access mechanisms to image data, concept-based image retrieval queries are expressed as keywords using either controlled or uncontrolled vocabularies supported by the retrieval system. In contrast, content-based image retrieval requires searchers to express their information problems as visual queries. *Query by visual example* (QVE) is the general term used to describe the overall approach and was first coined by Hirata and Kato (1992) to describe the interaction paradigm of their sketch query tool developed for the Art Museum image retrieval system developed by Kato, Kurita and Shimogaki (1989). Under this umbrella term, a number of query paradigms have been proposed for specifying a query Q_n in an image database I . Del Bimbo (1999) and Lew and Huang (2001) distinguish between different methods of QVE including query by browsing, query by icon, and query by colour. This research explores the usability of the query by image and query by sketch methods, which are outlined in Section 2.1.1 and Section 2.1.2.

To initiate a content-based image retrieval search, the user provides, selects or creates the visual representation of their information need. Smeulders et. al. (2000) defined an abstract query space to represent user interaction

in a content-based image retrieval system:

$$Q = \{I_Q, F_Q, S_Q, Z_Q\}$$

〈Equation 2〉 Abstract Query Space

where I_Q is the selection of images from the database I , F_Q is the selection of features derived from the images in I_Q , S_Q is the similarity function, and Z_Q is a set of labels to capture the goal dependent semantics. Rui and Huang(2001) proposed an alternative model to represent the query(Equation 3). They state that an image object O can be modeled as a function of the image data D , features F , and representations R :

$$O = O(D, F, R)$$

〈Equation 3〉 Image Object Model

where D is the raw image data, F is a set of visual features associated with the image object, such as colour, texture or shape, and R is a set of representations for a given feature such as colour histogram or colour moments. However, they suggest that specifying object O , feature F , and R representation level weights imposes a dual burden on the user of the retrieval system and its developers(〈Figure 2〉). For example, specifying the object level weights U imposes a burden

on the user, as it requires them to have comprehensive knowledge of the visual features used in the retrieval system and how they relate to their information need. Similarly, specifying features and their representation level weights V_i and W_{ij} imposes a burden on system developers who do not know which r_{ij} and r_{ijk} match users' perceptions of the image content.

In general, the majority of content-based image retrieval systems exemplify the interaction model proposed by Rui and Huang (2001). This approach is generally considered sufficient for the formulation of an initial query as it leads to a set of results which, it is suggested, can be iteratively refined to satisfy the searcher's information need. Yoshitaka and Ichikawa(1999) state that QVE provides the user with an intuitive method of query representation and a form of expressing a query condition that is close to that of the information problem. In contrast, Boujemaa, Fauqueur, and Gouet(2004) argue that while QVE has provided a framework for visual information retrieval research, it does not support the rich variety of visual search strategies required for effective image retrieval. However, there is little empirical evidence to support either of the polarised positions concerning the efficacy of QVE in supporting

the formulation of visual queries in a content-based image retrieval environment.

Eidenberger and Breitender(2002) argue that user interfaces for content-based image retrieval systems are overly complicated and that casual users would be overtaxed by the demands of selecting similarity measures, retrieval features, and setting of weights. They suggest that in order to improve the acceptance of future systems, simpler user interfaces are required, proposing a framework for their design as part of the VizIR project; an open-framework to develop a Java API for visual information retrieval. The framework includes a description of the visual components and their class structure, and the communications between panels, visual components and the query engines based on the Multimedia Retrieval Mark-up Language(MRML) proposed by Müller et. al.(2000). The proposed framework implements two methods of the QVE paradigm, which they consider are the most intuitive: query by image and query by sketch. While there is considerable merit in their proposal to develop a framework for user interfaces for visual information retrieval, their decision to implement the two query methods does not take into account whether these methods support the formulation of searcher's information problems.

2.1.1 Query by Image

Query by image allows the user to provide an example image as a representation of their query and is the most commonly implemented query method for content-based image retrieval. The example can be either an image selected externally or internally from the system. Di Lecce and Guerriero(1999) characterize this as query by external image example and query by internal image example. With query by external image example the searcher provides an external example image to the system. This method is generally perceived as the simplest approach to query formulation but it relies heavily on the end user having a suitable representative image to use as the basis of their query in the first instance. In contrast, query by internal image example allows the searcher select an image from the system's database and is a form of browsing. Typically, the user is presented with a randomly generated list in an n-dimensional matrix and then iteratively refines the retrieval results by selecting images that satisfy their information need. Systems that adopt this approach automatically compute a similarity ranking on image features in the set and disregard the least similar images after each iteration until the user is left with a set of relevant images(Laaksonen et. al. 2000; Vendrig,

Worring and Smeulders 2001).

Other approaches to query by internal image example include whether the data is unstructured or structured. With an unstructured approach, the searcher is presented with a complete view of all the images contained in the database and selects images that satisfy their information requirements. The view and structure of the image database can be modified using a relevance feedback mechanism. An early example of an unstructured browsing technique was proposed by Gecsei and Martin (1989). They suggested that users could progressively refine their search to one or two highly relevant surrogates with this approach. However, this was not demonstrated empirically. Similar approaches were proposed by Santini and Jain(1997) and Nakazato and Huang (2001) who extended this to 3D feature spaces including a 3D projection-based immersive VR using NCSA CAVE. However, they both acknowledge the difficulty of this approach in locating a suitable image to use as the basis of the initial query since no knowledge about the past or anticipated use of the system is available, the initial query space Q^0 should not be biased toward specific images or make some image pairs a priori more similar than others (Smeulders et. al. 2000). With a structured approach, images are clustered

into predefined, hierarchal classes based on a visual property or characteristic such as colour, shape or texture. The images are presented as thumbnails and the user can move through the hierarchy until satisfied. An example of this approach was proposed by Lai(2000) who clustered images into hierarchies using a set of ten colours based on an extended perceptual colour model proposed by Berlin and Kay(1991).

The concept of browsing has long been recognized in the field of information retrieval as a fundamental part of human information seeking behaviour(Wilson 1981). Marchionini(1995) states that browsing is an approach to information seeking that is both informal and opportunistic. Frost et al.(2000) suggest that one of the primary attractions of browsing is that it allows users to recognise what is interesting rather than formulating a precise information query in advance. Evidence suggests that on average users were willing to browse through approximately one hundred thumbnail images before refining the search parameters(2000). Nevertheless, the viability of browsing as an access strategy is heavily dependent on the size of the image collection and the effectiveness of any classification by which a physical order is imposed(1995).

2.1.2 Query by Sketch

Query by sketch allows the user to draw an approximate outline of the desired image with a graphic editing tool provided by the retrieval system. The sketch can represent either a completed object or scene. Kato et. al. (1989) were the first to provide a query by sketch interface for formulating visual queries but other examples of this approach have since been proposed by Chuah, Roth, and Kerpedjiev (1997), Muller, Eickeler, and Rigoll(1998), Di Sciascio and Mongiello(1999). Jaimes and Chang(2002) suggest that a sketch is a good way of making a precise request to a system because the user has total freedom in arranging different components when creating the query. In theory, any image could be retrieved using a sketch-based query as the user could potentially draw an exact replica of the image of interest. Nevertheless, Korfhage(1997) argued that queries formulated by this method are simplistic, relatively crude sketches of the desired query image and that the query method had limited functionality for expressing more complex image queries.

To address the limitations of the query by sketch method, Ko and Byun(2002) proposed query by gesture as an alternative. By this method, the user formulates a query by drawing the shape of the object using their finger

in front of a CCD(Charge-Coupled Device) camera, which captures finger motion at 160×120 image frames at a rate of five frames per second. After drawing a shape the user can then select colours from a palette and the number of k nearest neighbour retrieved results; the system also supported a modified probabilistic feature relevance feedback mechanism proposed by Ko, Peng, and Byun(2001) which allows the progressive refinement of query results. However, the results of a retrieval experiment using six queries against a database of three thousand images from the Corel stock photography library showed that average retrieval precision was approximately forty-five percent without relevance feedback which increased to sixty-four percent after five iterations. The results suggest that their proposed approach significantly affects the retrieval performance of the system, which is notably poor even after five iterations, and the results are far from encouraging.

Long, Zhand and Feng(2003) suggest that a coarse sketch is sufficient as the query can be refined based on retrieval results. However, it has been demonstrated that the effectiveness of shape matching features are highly sensitive to noise and pixel arrangement where shape similarity is computed by evaluating the correlation between a linear sketch and

edge images in the database(Del Bimbo and Pala 1997). As a result, the query image has an adverse effect on the retrieval results. Eakins(1992) argued that the shape drawn by the user must be an accurate representation of the shapes in the database, which he suggests is nearly impossible in practice although this was not demonstrated empirically.

At a theoretical level, query by sketch would appear to be a natural method for formulating visual queries. Despite its theoretical expressive power, the queries are difficult to formulate; drawing is a difficult task, particularly for users with limited skills. Blaser(1997) argues that the task can be simpler for users with some training or practice in drawing but, even then, using a computer interface to draw is usually unnatural. Regardless of whether the user is able to draw a good example to perform the query, the success of the system will depend on how the query is interpreted by the system. No research has been done to examine the effect of real user sketch-based queries on retrieval performance.

3. Experiment

In order to explore the viability of the query methods in supporting visual query for-

mulation a usability experiment was conducted. The aim of the experiment was to investigate to what extent query by image and query by sketch were effective methods for visual query formulation; the experiment was restricted to query by internal example. Retrieval performance was not measured or the focus of this study; the performance of the retrieval engine was investigated by Eakins, Boardman, and Graham(1998), which demonstrated that their approach, albeit for a prototype system under specific laboratory conditions, was capable of providing effective shape retrieval for the majority of potentially citeable images.

This experiment focused on usability as defined in the ISO 9241 standard: effectiveness, efficiency and satisfaction(ISO 1998). Effectiveness is the extent to which a goal or task can be achieved. Indicators of effectiveness include quality of solution and errors. Efficiency is the amount of effort required to complete a task. Indicators of efficiency include task completion time and learning time. Satisfaction is the level of comfort that the user feels when using a product or system and how effective that product or system was in supporting their goals. All three measures of usability were included in the experiment as Frøkjær, Hertzum, and Hornbæk(2000) argue that studies which focus solely on one or

two aspects risk making unfounded assertions about the overall usability of the system or product under investigation.

The sample population was drawn from the fourteen UK Patent Information Network (PIN) offices in order to produce a group who shared similar task orientation, motivation, system expertise, and domain knowledge. All were intermediary searchers acting on behalf of clients and who deal with the full range of trade mark information queries from non-specific to specific. The experiment included twenty-eight participants, four(14%) male and twenty-four(86%) female. The age range of participants was twenty-four to fifty eight ($\bar{x}=39.64$, $\sigma=9.639$) and had worked in the field of patent and trade mark information retrieval for periods ranging from less than one year to more than fifteen years. The participants had between one and fifteen years of experience of working with information systems with their level of information retrieval skill being described as either intermediate(50%) and expert(50%). All twenty eight participants had used a text-based image retrieval system and none had used a content-based image retrieval system prior to the experiment.

The experiment was a within-subjects experimental design. Participants were randomly

assigned, using a number generator, to condition A or condition B, where condition A was query by image and condition B was query by sketch to counterbalance learning sequence effects. The independent variable was query method. The dependent variables were task completion time, number of errors, and user satisfaction. The user satisfaction dependent also included the dependent variable system capabilities. For the purpose of the experiment, task completion time was defined as the total time taken to complete a task. The number of errors was defined as the total number of errors generated during a task. For the purpose of this experiment two error types were defined: E_1 and E_2 . A type E_1 error occurred when a participant could not take a task to completion. A type E_2 error occurred when an action did not lead to progress in performing the desired task.

Two questionnaires were designed to collect data on the sample population and their subjective opinions: a pre-session and post-session. The pre-session questionnaire collected demographic and background data. The post-session questionnaire was a modified version of the questionnaire for user interaction satisfaction (56) and was designed to ascertain the subjective satisfaction of participants. The granularity of the rating scale was modified

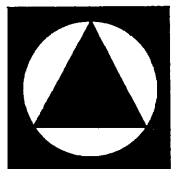
from its linear scale of 1-9 and NA to a negative and positive scale, as studies have suggested that this scale reading corresponds more accurately to the opinions of users(Hix and Hartson 1993). The scale contained no neutral value in order to reduce the error of central tendency thereby forcing participants' to make a negative or positive choice. Statistics were calculated for measures of central tendency, dispersion, and distribution. The data was analysed using the Wilcoxon signed-rank test at an alpha level of $p=0.001$.

Two separate sessions were held over consecutive days and each evaluation session lasted approximately one hour. The experiment consisted of six tasks per session: three benchmark and three intervening. Benchmark tasks are primary, representative tasks, which the system is specifically designed to support. Intervening representative tasks are secondary, inter-related tasks that users would be expected to perform, such as printing or saving search results. This class of task is not tested but is suggested as important in adding breadth and

depth to the evaluation(Preece et. al. 1994). The tasks were based on specific, semi-specific, and non-specific device mark image queries and took the form of paper-based examples, both final-copy or sketches, and descriptions of image queries. Benchmark tasks one T_1 was a specific query where the participant was presented with a final paper-based copy of a candidate mark image being registered(<Figure 1; Figure 2>).

Benchmark tasks two T_2 was a semi-specific query where the participant had examples of the overall shape and features of interest in the form of a paper-based sketch (<Figure 3; Figure 4>).

Benchmark tasks three T_3 was non-specific query where the participant had a general description of the overall shape or features of the device mark. The query by image task was to find images that contained 'wreath like shapes.' Similarly, the query by sketch task was to find an S shape that resembles a car racing track. The benchmark tasks were randomly selected, by a number generator,



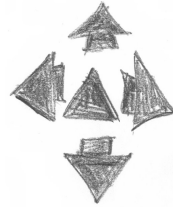
<Figure 1> Query by Image, Specific Task



<Figure 2> Query by Sketch, Specific Task



〈Figure 3〉 Query by Image,
Semi-Specific Task



〈Figure 4〉 Query by Sketch,
Semi-Specific Task

from fifteen hundred real queries submitted to the Patent Information Office and clustered by query type. The tasks were ordered from the specific to the non-specific to provide an increasing level of task difficulty. The number of tasks was restricted by the time constraints of the experiment.

4. Results

The following sections present a summary of the results of the experiment and are organized by the dependent variables of time, errors, and user satisfaction.

4.1 Time

Task completion time measured the total time taken to complete a task. <Table 1> shows the mean (\bar{x}), standard deviation (σ), variance (σ^2), and mean difference ($\bar{x}df$) for the dependent variable task completion time.

The results show that task T₁ and T₂ were completed in the shortest period of time using the query by image method with a mean difference of $\bar{x}df=3.7986$ and $\bar{x}df=4.4868$ between the two methods respectively. Task T₃ was completed in a shorter period of time using the query by sketch method with a mean difference of $\bar{x}df=4.42766$ between the two methods. The results also reveal that there was a notable degree of uniformity and con-

〈Table 1〉 Task Completion Time

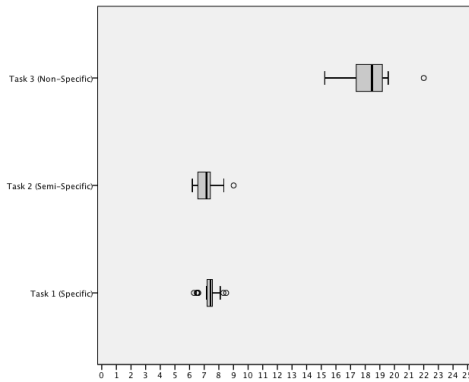
	Query by Image			Query by Sketch			
	\bar{x}	σ	σ^2	\bar{x}	σ	σ^2	$\bar{x}df$
T ₁	7.3643	0.52428	0.27487	11.1629	0.83627	0.69935	3.7986
T ₂	7.1400	0.66604	0.44361	11.6268	0.79987	0.63973	4.4868
T ₃	18.29696	1.38107	1.90736	13.8693	3.01110	9.06671	4.42766

sistency in overall task completion time scores for the query by sketch method with a marginal increase on task T₃. However, the variance in task completion time for T₃ using the query by sketch method reveals that the values are widely spread ($\sigma^2=9.06671$). In contrast, there was a marked increase in task completion time for task T₃ using the query by image method in comparison to the preceding two tasks with a mean difference of $\bar{x}df=11.15696$ between the highest and lowest scores. <Figure 5> and <Figure 6> show the distribution of scores for task completion time highlighting the significant increase in

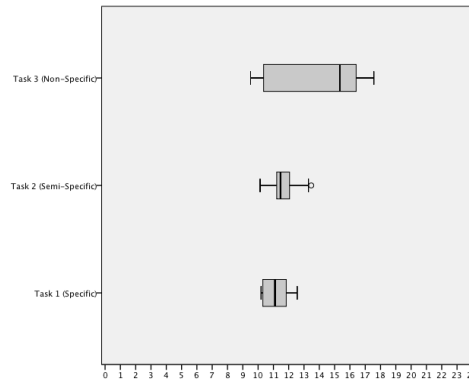
task completion time for task T₃ using the query by image method and the notable difference in the distribution in time for T₃ using the query by sketch method.

A wilcoxon signed rank test for individual task completion time reveals that there was a significant difference between the two query methods(<Table 2>).

<Table 3> shows the mean \bar{x} , standard deviation σ , and variance for the combined task completion time values of the query tools. <Figure 7> shows the distribution of data values for the combined scores. The results show that overall, tasks were completed in a shorter



<Figure 5> Query by Image, Task Time



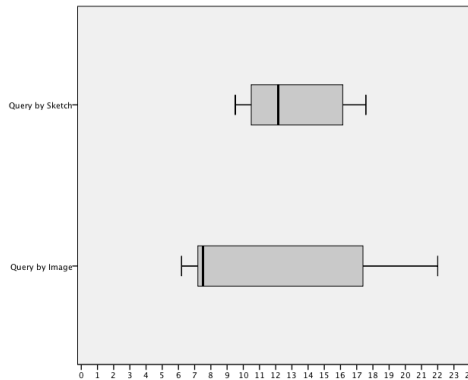
<Figure 6> Query by Sketch, Task Time

<Table 2> Wilcoxon Signed Rank Scores for Task Completion Time by Task Type

	Wilcoxon Signed Rank
T ₁	Z = -4.623, p < 0.001
T ₂	Z = -4.623, p < 0.001
T ₃	Z = -4.444, p < 0.001

〈Table 3〉 Combined Task Completion Time Scores

	\bar{x}	σ	σ^2
Query by Image	10.9246	5.30681	28.162
Query by Sketch	13.1218	2.69034	7.238



〈Figure 7〉 Distribution of Combined Task Completion Scores

time using the query by image method. However, the variance score for the query by image method shows the task completion time was much more widely distributed in comparison to the query by sketch method when the data is taken as a whole.

The results of a wilcoxon signed-rank test for the combined task completion time scores revealed a significant difference between query by image and query by shape $Z=-3.597$, $p<0.001$.

4.2 Errors

Two error types were defined for this experiment: E_1 and E_2 . A type E_1 error occurred

when a participant could not take a task to completion. A type E_2 error occurred when an action did not lead to progress in performing the desired task. In this experiment, no E_2 type errors were produced by the sample population. In contrast, eleven E_1 type errors occurred on task T_3 using the query by sketch method, which was only completed successfully by seventeen out of the twenty-eight participants with the other eleven participants abandoning the task. The results of this are discussed in the next section. Tasks T_1 and T_2 were all successfully completed by the twenty-eight participants using both the query methods. A wilcoxon signed rank test for E_2 type errors on individual tasks reveals that

<Table 4> Wilcoxon Signed Rank Scores for E2 Type Errors

	Wilcoxon Signed Rank
T ₁	Z=0.000, p=1.000
T ₂	Z=0.000, p=1.000
T ₃	Z=-3.317, p<0.001

there was a significant difference between the two query methods on task T₃, Z=-3.317, p<0.001 (<Table 4>).

4.3 Satisfaction

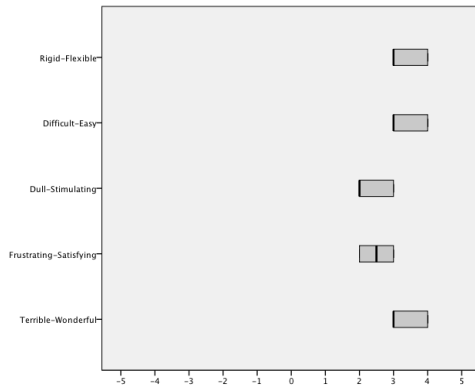
<Table 5> shows the mean \bar{x} and standard deviation σ for overall reaction to the query methods. The results reveal that the query by image method was rated positively on all five measures. Query by image was considered wonderful (\bar{x} =3.43), satisfying (\bar{x} =2.50), stimulating (\bar{x} =2.29), easy (\bar{x} =3.39) and flexible (\bar{x} =3.34). In contrast, the query by sketch method was rated negatively on four out of the five semantic differentials. The method was considered, terrible (\bar{x} =-1.68), frustrating (\bar{x} =-2.61), dull (\bar{x} =-1.32), and

difficult (\bar{x} =-3.46). However, the method was rated positively on the rigid-flexible semantic differential (\bar{x} =1.25). The difference between the scores for the two query methods on the difficult-easy semantic differential is diametrically opposed with the query by image method being rated highly positively (\bar{x} =3.39) and the query by sketch method being rated highly negatively (\bar{x} =-3.46). In terms of overall reaction, this data suggests that this sample population considers the query by sketch method insufficient for visual query expression and its ease of use is highly questionable.

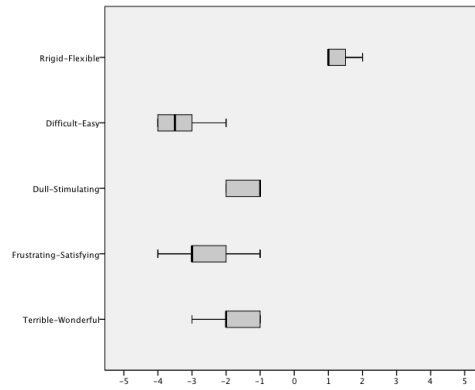
<Figure 8> and <Figure 9> show the distribution of scores for overall users reaction to the query methods. The results reveal that there is a clustering between the individual

<Table 5> Overall Satisfaction, Mean and Standard Deviation

	Query by Image	Query by Sketch
Terrible v Wonderful	3.43 (0.504)	-1.68 (0.670)
Frustrating v Satisfying	2.50 (0.509)	-2.61 (0.737)
Dull v Stimulating	2.29 (0.460)	-1.32 (0.476)
Difficult v Easy	3.39 (0.497)	-3.46 (0.576)
Rigid v Flexible	3.43 (0.504)	1.25 (0.441)



〈Figure 8〉 Distribution Scores, Query by Image



〈Figure 9〉 Distribution Scores, Query by Sketch

scores within range of one or two points in the inter-quartile range and that the scores are not widely distributed which suggests a degree of consensus between participants' ratings.

A wilcoxon signed rank test on individual measures shows that there was a significant difference, $p < 0.001$, in user satisfaction between the two tools <Table 6>.

<Table 7> shows the mean \bar{x} , standard deviation σ , and variance σ^2 for overall reaction to the query methods. The results reveal that the query by image method.

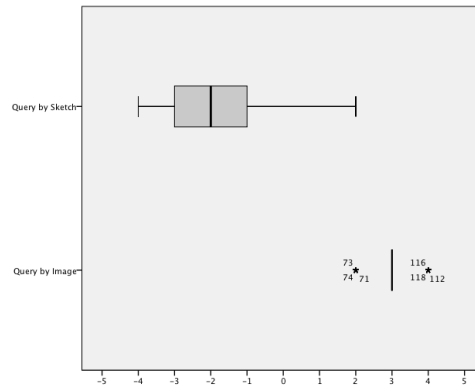
<Figure 10> shows the distribution of scores for overall user reaction. The results show that there is a tight clustering in the scores for overall satisfaction with the query

〈Table 6〉 Wilcoxon Signed Rank Scores for Individual Measures of Satisfaction

	Wilcoxon Signed Rank
Terrible v Wonderful	$Z = -4.692, p = 0.001$
Frustrating v Satisfying	$Z = -4.676, p = 0.001$
Dull v Stimulating	$Z = -4.736, p = 0.001$
Difficult v Easy	$Z = -4.697, p = 0.001$
Rigid v Flexible	$Z = -4.878, p = 0.001$

〈Table 7〉 Combined Scores for Overall User Satisfaction

	\bar{x}	σ	σ^2
Query by Image	3.00	0.705	0.496
Query by Sketch	-1.56	1.702	2.895



〈Figure 10〉 Distribution of Combined Scores for Overall User Satisfaction

〈Table 8〉 Efficacy of the Query Methods

	\bar{x}	σ	σ^2
Query by Image	2.61	0.567	0.321
Query by Sketch	-3.32	0.772	0.597

by image method suggesting a strong degree of consensus in participants' ratings. Similarly, the distribution of scores for the query by sketch method is within range of one or two points in the inter-quartile range and that the scores are not widely distributed.

The results of a wilcoxon signed-rank test for the combined satisfaction scores revealed a significant difference between query by image and query by shape $Z=-10.223$, $p<0.001$.

In addition to overall satisfaction, participants were also asked to consider the efficacy of the query methods. <Table 8> shows the mean \bar{x} , standard deviation σ , and variance for participants perceived usefulness of the query tools. The results show that query

by image was rated positively ($\bar{x}=2.61$). In contrast, query by sketch was rated highly negatively ($\bar{x}=-3.32$).

The results of a wilcoxon signed-rank test for the efficacy scores revealed a significant difference between query by image and query by shape $Z=-4.672$, $p<0.001$.

5. Discussion

Query by image performed well across all three measures of usability. From an effectiveness perspective, task completion time scores reveal that specific and semi-specific tasks were completed in a shorter time using

this method. However, there was a marked increase for non-specific tasks. This increase in task completion time can be attributed specifically to the level of task difficulty. As the results of the individual task completion time scores <Table 1> show, there was a marked difference between specific and non-specific tasks. This is not a surprising result in itself as the tasks were specifically designed to provide an increasing level of complexity. As a result, it could be predicted that it would take participants longer to complete tasks that had a higher cognitive load as a result of the increased level of subjective interpretation introduced by the task. Whether this can be solely attributed to the complexity of the task and the increasing level of cognitive processing required by the searcher to solve the information problem or whether it is related to limitations of the query method itself is open to discussion.

Query by sketch is a common approach to visual query formulation. However, there is considerable disagreement between researchers as to the efficacy of this method in supporting visual query formulation. It is broadly considered either as an aid or a hindrance to effective and efficient query formulation. Despite being one of the main query methods used in content-based image retrieval proto-

type systems the results of this study suggest that the query by sketch method does not support visual query formulation. Examination of the indicators of usability shows extremely disappointing performance measures. In terms of satisfaction, participants overall satisfaction ratings for query by sketch were considerably less favourable across all five high-level user interface factors in comparison to query by image. The method was only rated positively on the 'rigid v flexible' semantic differential, which suggests that participants still considered that the query method offered some perceived degree of flexibility, at least from a theoretical perspective. However, the utility of the tool to aid visual query formulation is clearly not reflected in their ratings scores for other measures of satisfaction.

Despite its perceived expressive power, it was observed that even specific queries were extremely difficult to formulate. Drawing can be a difficult and challenging task particularly for people with limited natural ability. It was observed that queries formulated by this method were relatively crude, poorly formulated sketches, even when the task showed the exact query image. This supports the position taken by Korfhage(1997) who argued that query by sketch lacked the functionality to express an information problem beyond

simple shapes. To what extent this may also be exacerbated by the input device is unclear and a comparison of different methods of sketching or drawing with a mouse and a graphics tablet using pencil and paper as a control would prove insightful. Whether this can be overcome with training and practice as Blaser(1997) suggests is unclear and provides an opportunity for further research to investigate the limits of the approach in supporting visual query formulation. In the context of this study, it was observed that participants found the task of drawing by this method challenging and frustrating. It was also observed that the quality of the query, particularly in terms of its structure, had a notable effect on the retrieval results. This is contrary to this position taken by Long, Zhand and Feng(2003) who suggested that a rough sketch would prove sufficient to produce a set of results, which could be refined. Similarly, while the results of retrieval performance experiments suggest that the retrieval effectiveness of existing prototype systems is extremely encouraging under specific laboratory conditions they also suggest that the quality of the query has an effect on retrieval performance. To what extent it affects the retrieval result is unclear as little empirical evidence exists into the effect of query formulation and its ef-

fect on retrieval performance.

For this experiment two error conditions were defined. Query by sketch produced a number of E1 type errors during the experiment where participants could not take a task to completion. This task involved using the query method to formulate a non-specific query with eleven participants abandoning the task. In comparison to the findings of Bird, Elliott, and Hayward(1999), it was noted that participants demonstrated considerable frustration in their attempts to produce a query of sufficient quality and the effect that this had on the retrieval results, blaming their ability to use the tool correctly. Given the expertise of the sample population with information systems and user interfaces, it is interesting to note that they attribute their ability as the single point of failure in the system rather than the usability of the system itself. This observation is not uncommon where users attribute usability problems experienced with a system to their ability to use it effectively. In addition, an error classification with a finer level of granularity, such as that proposed by Norman(1981), may have revealed more interesting data regarding the type of errors experienced by participants during the course of the experiment. Nevertheless, the development of such taxonomy was beyond the scope of

this research.

These findings have important implications for the efficacy and utility of content-based image retrieval as an approach although there is a need to further examine the usefulness of both the tools in relation to retrieval performance which will be the focus of future research work.

6. Conclusion

This paper investigated the efficacy of the query by image and query by sketch methods in supporting a range of information problems through a usability experiment in order to contribute to the gap in knowledge regarding the relationship between searchers' information problems and the query methods required to support efficient and effective visual query formulation. The results of the experiment suggest that query by image is a viable approach to visual query formulation. In contrast, the results strongly suggest that there is a significant mismatch between the searchers information problems and the expressive power of the query by sketch paradigm in supporting visual query formulation.

Content-based retrieval research has generally operated in a vacuum isolated from the

real activities and tasks of the searcher focusing primarily on the technical aspects of delivering effective image retrieval. As a consequence, the design of query interfaces for content-based image retrieval has been driven by the underlying retrieval mechanism and not by the requirements of the end user. The importance of the user in the design and the evaluation of content-based image systems have been grossly underestimated by the research community. What the information needs of users of image collections are and how successfully these can be expressed in a content-based image retrieval environment is far from clear. This is a complex field and no single solution or approach will solve the multi-dimensional issues to achieve effective and efficient retrieval of the data. The fusion of work from a diverse range of disciplines, including mathematics and statistics, information and computer science, neuroscience and psychology, will all contribute to the development of accessible image retrieval systems. Nevertheless, research into query interfaces for content-based image retrieval is still in its infancy and substantial research is required to understand the relationship between the user's information problem and methods of query formulation and refinement.

References

- Berlin, B. and P. Kay. 1991. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
- Bird, C., P. J. Elloit, and P. M. Hayward. 1999. "Content-Based Retrieval for European Image Libraries." In: *The Challenge of Image Retrieval: CIR99*, The 2nd UK Conference on Image Retrieval, Forte Posthouse Hotel, Newcastle upon Tyne, United Kingdom, 25th-26th February 1999, edited by D. J. Harper and J. P. Eakins.
- Blaser, A. D. 1997. "User Interaction in a Sketch-Based GIS User Interface." *Lecture Notes In Computer Science*, 1329 (1997): 505-506.
- Bordogna, G. and M. Pagani. 2010. "A Flexible Content-based Image Retrieval Model and Customizable System for the Retrieval of Shapes." *Journal of the American Society for Information Science and Technology*, 61(5): 907-926.
- Boujemaa, N., J. Fauqueur, and V. Gouet. 2004. "What's Beyond Query by Example?" *Lecture Notes in Computer Science*, Springer-Verlag.
- Castelli, V. and L. D. Bergman. 2002. *Image Databases: Search and Retrieval of Digital Imagery*. New York: John Wiley, 1-10.
- Chin, J., V. Diehl, and L. Norman. 1988. "Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface." In: *CHI '88: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Washington, D.C. USA, 15-19 May 1988, edited by J. J. O'Hare, ACM Press: 213-218.
- Datta, R., D. Joshi, J. Li, and J. Z. Wang. 2008. "Image Retrieval: Ideas, Influences, and Trends of the New Age." *ACM Computing Surveys*, 40(2).
- Del Bimbo, A. and P. Pala. 1997. "Visual Image Retrieval by Elastic Matching of User Sketches." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2): 121-132.
- Di Lecce, V and A. Guerriero. 1999. "An Evaluation of the Effectiveness of Image Features for Image Retrieval." *Journal of Visual Communication and Image Representation*, 10(4): 351-362.
- Eakins, J. P. 1992. "Pictorial Information Systems: Prospects and Problems." *Proceedings of 14th BCS Information Retrieval Specialist Group Research Colloquium*, Lancaster, April 1992:

- 102-123.
- Eakins, J. P., J. M. Boardman, and M. E. Graham. 1998. "Similarity Retrieval of Trademark Images." *IEEE MultiMedia*, 5(2): 53-63.
- Eidenberger, H. and C. Breiteneder. 2003. "VizIR: A Framework for Visual Information Retrieval." *Journal of Visual Languages and Computing*, 14(5): 443-469.
- Enser, P. G. B. 1995. "Progress in Documentation Pictorial Information Retrieval." *Journal of Documentation*, 51(2): 126-170.
- Feng, D., W. C. Siu, and H. J. Zhang. 2003. *Multimedia Information Retrieval and Management: Technological Fundamentals and Applications*. New York: Springer, 1-26.
- Flickner, M., H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. 1995. "Query by Image and Video Content: The QBIC System." *IEEE Computer*, 28(9): 23-32.
- Frøkjær, E., M. Hertzum, and K. Hornbæk. 2000. "Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?" *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, The Hague, The Netherlands, 01-06, April, 2000, ACM Press: 345-352.
- Gecsei, J. and D. Martin. 1989. "Browsing Assess to Visual Information." *Optical Information Systems*, 9(5): 237-241.
- Gudivada, V. N. and V. V. Raghaven. 1995. "Content-based Image Retrieval Systems." *IEEE Computer*, September 1995: 20.
- Hanjalic, A., N. Sebe, and E. Y. Chang. 2006. "Multimedia Content Analysis, Management and Retrieval: Trends and Challenges." In: *Electronic Imaging 2006, Multimedia Content Analysis, Management and Retrieval 2006*, IS&T/SPIE International Symposium, SPIE Vol. 6073, San Jose Marriott and San Jose Convention Centre, San Jose, California, USA, 15-19 January 2006, edited by Edward Y. Chang, Alan Hanjalic and Nicu Sebe, SPIE.
- Hirata, K. and T. Kato. 1992. "Query by Visual Example." *Proceedings of the 3rd International Conference on Extending Database Technology*, Vienna, Austria, 23rd-27th March 1992, edited by A. Pirotte, C. Delobel and G. Gottlob, Springer-Verlag: 56-71.
- Hix, D. and H. R. Hartson. 1993. *Developing User Interfaces: Ensuring Usability Through*

- Product and Process*. New York: John Wiley.
- Idris, F. and S. Panchanathan. 1997. "Review of Image and Video Indexing Techniques." *Journal of Visual Communication and Image Representation*, 8(2): 146-166.
- ISO 9241. 1998. International Organization for Standardization.
- Jaimes, A. and S-F. Chang. 2002. "Concepts and Techniques for Indexing Visual Concepts." In: Castelli, V. and Bergman, L. D. (ed). *Image Databases: Search and Retrieval of Digital Imagery*, John Wiley & Sons Inc. : 497.
- Jaimes, A., N. Sebe, and D. Gatica-Perez. 2006. "Human-centered Computing: A Multimedia Perspective." *Proceedings of the 14th annual ACM international conference on Multimedia*. Santa Barbara, CA, USA, October 23-27, 2006, ACM: 855-864.
- Kato, T. 1991. "Intelligent Visual Interaction with Image Database Systems: Toward the Multimedia Personal Interface." *Journal of Information Processing*, 14(2): 134-143.
- Kato, T., T. Kurita, and H. Shimogaki. 1989. "Multimedia Interaction with Image Database Systems." In: *Advanced Database System Symposium '89*, Kyoto, Japan, 1989: 271-278.
- Kimia, B. B. 2002. "Shape Representation for Image Retrieval." In: Castelli, V. and Bergman, L. D. (eds). *Image Databases: Search and Retrieval of Digital Imagery*, John Wiley & Sons Inc. : 345-372.
- Ko, B. C. and H. Byun. 2002. "Query-by-Gesture: An Alternative Content-Based Image Retrieval Query Scheme." *Journal of Visual Languages and Computing*, 13(4): 375-390.
- Ko, B. C., J. Peng, and H. Byun. 2001. "Region-Based Image Retrieval Using Probabilistic Feature Relevance Learning." *Pattern Analysis and Applications*, 4(2-3): 174-184.
- Korfhage, R. R. 1997. *Information Storage and Retrieval*. New York: John Wiley.
- Laaksonen, J., M. Koskela, S. Laakso, and E. Oja. 2000. "PicSOM: Content-based Image Retrieval with Self-organizing maps." *Pattern Recognition Letters*, 21(3-14): 1199-1207.
- Lai, T-S. 2000. "CHROMA: A Photographic Image Retrieval System." Ph.D. Thesis. University of Sunderland, UK. January 2000.
- Lew, M. S. and T. S. Huang. 2001. "Visual Information Retrieval: Paradigms, Applications, and Research Issues." In: Lew, M. S. (ed.) *Principles of Visual Information Retrieval*. Springer-Verlag, Ad-

- vances in Pattern Recognition: 3-9.
- Loncaric, S. 1998. "A Survey of Shape Analysis Techniques." *Pattern Recognition*, 31(8): 983-1001.
- Preece, J., Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey. 1994. *Human-computer Interaction*. Boston: Pearson Addison Wesley.
- Mingqiang, Y., K. Kidiyo, and R. Joseph. 2008. *A Survey of Shape Feature Extraction Techniques*. Pattern Recognition Techniques, Technology and Applications, P-Y. Yin (Ed.), InTech: 43-90.
- Müller, W., H. Müller, S. Marchand-Maillet, T. Pun, D. M. Squire, Z. Pecenovic, C. Giess, and A. P. de Vries. 2000. "MRML: An Extensible Communication Protocol for Interoperability and Bench-marking of Multimedia Information Retrieval Systems." *Visual Information and Information Systems*: 300-311.
- Nakazato M. and T. S. Huang. 2001. "An Interactive 3d Visualization for Content-based Image Retrieval." *Proceedings of the International Conference on Multimedia and Expo, Waseda University, Tokyo, Japan, 22-25 August 2001*: 44-47.
- Rui, Y. and T. S. Huang. 2001. "Relevance Feedback Techniques." In: Lew, M. S. (ed.) *Principles of Visual Information Retrieval*. Springer-Verlag, Advances in Pattern Recognition: 219-258.
- Santini, S. and R. Jain. 1997. "Image Databases are not Databases with Images." *Proceedings of the 9th International Conference on Image Analysis and Processing*. Volume II: 38-45.
- Smeulders, A. W. M., M. Worring, S. Santini, A. Gupta, and R. Jain. 2000. "Content-based Image Retrieval at the End of the Early Years." *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(12): 1349-1380.
- Veltkamp, R. C. and M. Hagedoorn. 2001. "State of the Art in Shape Matching." In: Lew, M. S. (ed.) *Principles of Visual Information Retrieval*. Springer-Verlag, Advances in Pattern Recognition: 87-120.
- Vendrig, J., M. Worring, and A. W. M. Smeulders. 2001. "Filter Image Browsing: Interactive Image Retrieval by Using Database Overviews." *Multimedia Tools and Applications*, 15(1): 83-103.
- Wilson, T. D. 1981. "On User Studies and Information Needs." *Journal of Documentation*, 37: 3-15.
- Yoshitaka, A. and T. Ichikawa. 1999. "A Survey on Content-Based Retrieval for Multimedia Databases." *IEEE Transactions on Knowledge and Data Engineering*, 11(1): 81-93.