

画像検索のための3D インターフェースとシステム適応検討

小池真由美*1 青木輝勝*2 池田佳代*3 伊藤学*2 日高宗一郎*4

*1 (有) エスパリエ *2 東京大学先端科学技術研究センター
*3 (有) エクセリードテクノロジー *4 国立情報学研究所

近年、インターネット等のネットワークを用いたコンテンツ流通が盛んに行われるようになり、画像検索に関するニーズが急速に高まってきているが、現在画像検索技術は必ずしも実用レベルに達しているとは言えず、今後さらなる改良に向けた研究開発が必要である。この画像検索の課題が入力インターフェースにあることを踏まえ、本稿では「検索したい画像」内の対象物体が多面体の場合、対象物体の撮影されている面数に比例して従来の2D インターフェースでの描画が困難になることを実証し、筆者らの提案してきた3D インターフェースの有効性を示す。

3D Interface for Image/Video Retrieval and A Study on Its Application to Retrieval System

Mayumi Koike*1 Terumasa Aoki*2 Kayo Ikeda*3
Manabu Ito*2 Soichiro Hidaka*4

*1 Espalier Inc. *2 University of Tokyo, RCAST
*3 Excellead Technology
*4 National Institute of Informatics

In recent years, content distribution through the Internet has been actively carried out, and image/video distribution as well as music distribution has become widespread in rapid pace. Under this situation, there are growing needs for image/video retrieval, however, technologies for image/video retrieval have not reached to practical level. As such, more research will be required. Based on the assumption that the biggest issue in current image/video retrieval is in input interface, the paper describes the effectiveness of 3D input interface, which has been proposed by authors, from the result of verification test showing that as the number of face taken in the "target picture" increases, difficulties in drawing the picture with 2D interface increase.

1. はじめに

近年、個人の画像の収集量は、かつてないスピードで増加し続け、そのデジタル化も進んでおり、個人所有のコンテンツやインターネット上の画像から所望シーンを精度よく検索するユーザ要求は非常に強まっている。

しかし、依然として個人が使いやすい形での画像検索技術は実用化されているとはいえない。これは、検索目標とその目標画像に対する記憶が曖昧であり、目標画像のイメージを正確に描けない（表現できない）ことに起因している。従って、高精度な画像検索の実現のためには、ユーザの検索要求をどのように入力し、システムがその入力をどのように解釈・処理し、結果をユーザに返すかというユーザインターフェースの観点からの検討が必要であり、特にユーザのクエリ（Query）生成をどのように支援するかが画像検索技術における最大の課題であると言える。

2. 3D インターフェースの提案

2.1 従来の検索インターフェース

インデキシングコストを抑えつつ高精度な検索を行うには、テキスト語句の厳密性と略画の曖昧さを組み合わせたクエリ入力インターフェース、即ち、定義が明確な名詞（固有名詞等）と曖昧な記憶ながら概略だけは描ける略画の2つの入力を組み合わせたインターフェースが最も優れていることが容易に推測される[1]が、本稿では、このうちの略画入力の方に焦点を当てる。

「絵を描く」という行為はたとえ簡易略画とは言え一般的には一部の人間を除くと非常に苦手としているのが現状である。この要因としては、

- (1) 正確な形状・色の再現が困難であること。
- (2) 物体とは元来すべて3Dであるのに2Dで表現しなければならないこと。

の2点が挙げられる。そこで筆者らは特に(2)の負担を軽減することで、より直感的なインターフェースが実現することを目的に3D入力インターフェースを提案している[1]。

2.2 3D入力インターフェースの可能性

3D入力インターフェース*とは一般の静止画像/動画を検索する際に3D情報を入力クエリと

して用いる手法である。一例として図1を「検索したい画像」として取り上げ、あらかじめこの画像を見たことがある検索者が、この写真を探す場面を想定することにする。

この場合、略画を描くためにはこの被験者の頭の中では2D的な情報、すなわち、「左右にりんごがあり、左のりんごは右のりんごと比べて高さ、幅がおおよそ半分ずつである。またこの2つはほとんどくっつきそうなくらい近づいている。」という記憶がなされていなければ正確な略画を描くことはできないことになる。しかし人間は本来このような記憶法をとっていないことは殆ど自明であろう。



図1 検索したい画像

一方、この写真を見たときに一般的には人間は2D写真でありながら3D的な情報、例えば「手前のりんごと奥のりんごは同じくらい大きさで奥行き方向に50cmくらい離れている」などの見方で脳に記憶させている。

このため、より直感的な入力インターフェースを設計するにあたっては3Dモデリング的手法を取り入れることが望ましい。

これより、筆者らは「3D入力インターフェース」として、図2のような方式を提案している。

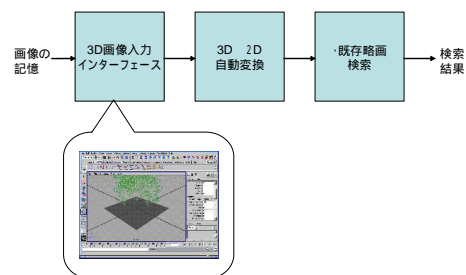


図2 3D入力インターフェース

一般的な市販3DCG制作ソフトウェアのような

インターフェースを有し、直方体、立方体、球、回転体などが極めて容易に作成できるモジュールを用い、そこから、カメラワークの各種技法（クローズアップショット、ウェストショット、など15種類程度）を用いてそれぞれ並行的に作成した3Dモデルを”撮影”をするものである。これはまさに3Dモデリングを2D略画に変換する処理に他ならない。最後にこのように撮影された15枚程度の略画を従来同様の略画検索ツール[2]に入力させ、最終的な演算結果（検索結果画像）を得る。

2.3 3D入力インターフェースのための予備実験結果

2章で述べた如く、物体は元来3Dであることから、従来の2D入力インターフェース入力での課題は、奥行きと位置関係の表現であろう。

そのため、『3Dでの入力が直感的なものとなりえるか』、つまり、奥行きと位置関係を2Dインターフェースに比してより正確に表現できるか、を筆者らは実験を通じて検証してきた。

過去の筆者らによる、実験[3][4][5]においては、以下の3点が実証されている。

- (1) 2D描画での、2個の物体の相関関係は、横より奥行きを把握することが難しい。
- (2) 3D配置の場合、2D描画より2個の物体の奥行きの相関関係が把握しやすい。
- (3) 奥行きの把握しやすい3D配置において、白と黒の球での場合、白の球がより近く、黒の球がより遠く認識・表現される。

3. 3D 入力インターフェースの画像

検索システムへの適用

3.1 入力インターフェース

2.3章で述べたように、画像検索システムにおいて3D入力インターフェースが2D入力インターフェースより高い検索効率を実現する可能性が示されたが、この有効性を実現する3D入力インターフェースのためには、図2で図示した最初のステップ、すなわち「3D画像インターフェース」において、以下の2つの要件があると筆者らは仮定する。

- (1) 2D画像インターフェースと同様の使いやすさ・精度・効率を実現できること

- (2) カメラワークを使用できることである。

(1)は、たとえ3D（奥行き・角度）表現が不要な画像でも簡単に略画がかけられるために必要であり、(2)は(1)で実現した略画の3D化（奥行き・角度）表現に必要なためである。既存のモデリングソフトを検索用の3D入力インターフェースと考えた場合、(1)と(2)は両立しないケースが多い。3D描画ソフトは、簡易なものは(1)に重点を置き(2)がなく、プロ向けのものは(2)は存在するが、(1)が実現しない。

(1)で簡単に3D表現ができれば、(2)は不要である。そこで、筆者らは、「検索したい画像」内の対象物体が多面体の場合、対象物体の撮影角度による撮影される面数と、ユーザの描画の困難性との関係を調べる実験を行った。一般的ユーザはソフトウェア上の、2D入力インターフェースと3D入力インターフェースでは前者の経験が豊富であろうと仮定し、その2D描画で3種の絵を「検索したい画像」として描画をしてもらった。

3.3 実験内容

本稿で実施した実験の概要を示す。

- ・使用ソフトウェア：Microsoft Powerpoint
- ・実験使用画像枚数：3枚（図3）
ブルーバックにオレンジの直方体を以下の位置から撮影したもの。
 - 水平位置から
 - 上部から
 - 上部正面左側から
- ・被験者数：5名

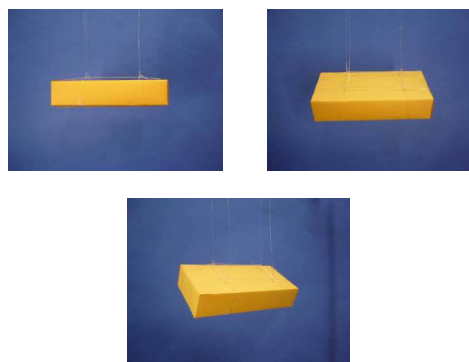


図3 実験使用画像

- ・実験手順と方法：

(1)被験者に「検索したい画像」として実験使用画像を見せる（Power Point 上にて）

(2)被験者が「検索したい画像」を Power Point 上にて描画する。

(3)(1)(2)にかかる時間を計測する。

(4)(1)-(3)を実験対象画像 - の各々で繰り返す。

(5)(2)の描画を「検索したい画像」と比較する。

・被験者へのアンケート

実験に際して、被験者には、以下の「描画に関するアンケート」に5段階（日常的に描く 時々描く たまに描く ほとんどない ない）に回答をもらっている。

(1) 紙の上で何かの絵を描いたことがある

(2) コンピュータ上で何かの絵を描いたことがある

(3) パワーポイントで何かの絵を描いたことがある

(4) 3Dソフトウェアを使って何かの絵を描いたことがある

(5) コンピュータ上で何かの絵を描いて画像を検索したことがある

3.5 実験結果

3.4において設定した実験を行ったところ、以下の結果が導かれた。まず、アンケートに対する回答を表1に示す。

回答 質問番号	日常的に	時々	たまに	殆どない	ない
(1)			2	3	
(2)		3		2	
(3)	3			1	1
(4)				2	3
(5)				2	3

表1 アンケート結果

これにより、被験者は3D描画を行った経験はまれであり、描画による検索を行った経験も稀であることがわかる。パワーポイントでの2D描画は、日常的な経験者と未経験者に分かれた。

次に、実験において描画にかかった時間について表2示す。

(単位:秒)

撮影面数	被験者1	被験者2	被験者3	被験者4	被験者5
1	50	100	67	25	33
2	115	225	145	65	72
3	230	510	322	140	155

表2 撮影面数と描画時間

これにより、「検索したい画像」に映る多面体の面数が増えるほど描画時間がかかることがわかる。

次に、3.3節で示した各画像の描画結果を図3、4、5にそれぞれ示す。(図内 - は、被験者番号)

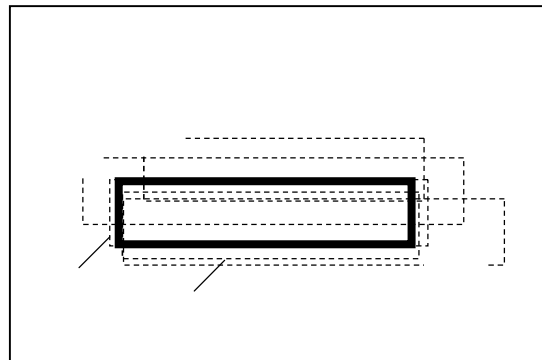


図3 「検索したい画像」の描画結果

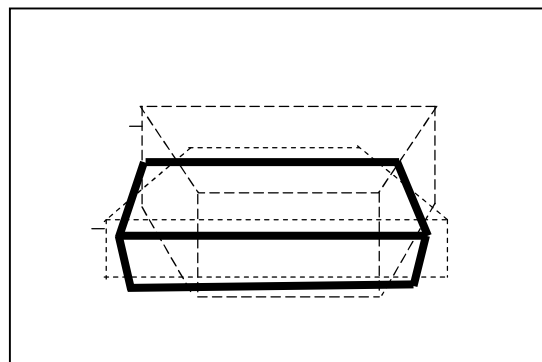


図4 「検索したい画像」の描画結果(一部)

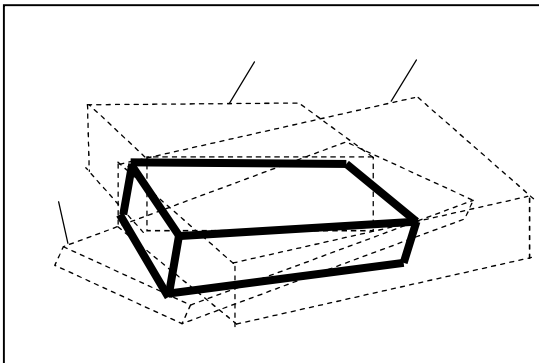


図5 「検索したい画像」の描画結果(一部)

「検索したい画像」については、多少の高さ・長さが異なるものの、被験者全員が「検索したい画像」に近い描画ができるのに対し、「検索したい画像」、と、面数が増加するに従い、

- ・対象形状
- ・でほぼ把握されていた高さや長さの比率

にも誤りが生じる結果となった。

4.まとめ

本稿では、3D入力インターフェースの検索システム適用に際し、

- ・3D入力インターフェース部の仕様検討
- ・2Dインターフェースにおいては多面体の面数の増加に従って描画の時間と困難性が増すことの実証

を行った。

3Dインターフェースを使用すれば、多面体の回転もしくは、カメラワークの使用によって、描画の困難性を緩和することが可能となる。

検索目的の描画における最終目的は、時間をかけて精度よい描画を完成させることではなく、検索を効率よく行うことなので、3.1節で述べた2要件を満たすソフトウェアが実現できれば、2.3節の3点の予備実験結果をソフトウェアでも検証することの第一歩となる。それができれば3Dインターフェースの有効性を確認する実験に使用できる「検索したい画像」の対象は大幅に広がる。例えば予備実験と同様に2物体を対象にした実験だけでも、物体色、背景色の2要素を簡単に変更でき

るため、既知である、色相が与える奥行き感(暖色系は進出し、寒色系は後退して見える)[8]と、手書き検索との関係についてデータが取得でき、3D入力インターフェースへの適用に役立つであろう。

今後の課題としては、検索適用の3D入力インターフェースを設計・構築し、そこへ描画・入力したユーザ入力画像を実際の検索システムに適用し、これ迄実験で検証してきた3D入力インターフェースの有効性が実現可能かの実験を実施する。

また、「検索したい画像」の対象を広げ、画像内の描画対象物体の形状、色、配置、個数を変更した場合、3m以上の距離で撮影された画像の場合、等の実験も必要である。

謝辞：本研究は総務省戦略的情報通信研究開発推進精度研究主体育成型研究開発平成15年度「簡単映像コンテンツ制作のための高度映像検索技術に関する研究(研究開発)」(研究代表者:青木輝勝(東京大学))の一環として行われたものである。尚、本実験に参加いただいた7人の被験者に心より感謝申し上げます。

文献

- [1] 小池真由美, 青木輝勝, 池田佳代, 伊藤学, 日高宗一郎「画像検索のための3Dインターフェース」情報処理学会 オーディオビジュアル複合情報処理研究会, 2004.3
- [2] 青木秀一, 青木輝勝, 安田浩, "動画からのシーン検索のための略画処理手法の提案", 情報処理学会 CVIM 研究会, 2002.1
- [3] 小池真由美, 青木輝勝, 池田佳代, 伊藤学, 日高宗一郎「画像検索のための3Dインターフェースの有効性の検証」情報処理学会 オーディオビジュアル複合情報処理研究会, 2004.6
- [4] 小池真由美, 青木輝勝, 池田佳代, 伊藤学, 日高宗一郎「高度画像検索のための直感的インターフェース」FIT2004, 2004.9
- [5] 小池真由美, 青木輝勝, 池田佳代, 伊藤学, 日高宗一郎「画像検索のための3D入力インターフェースの有効性検証実験」画像電子学会第213回研究会, 2004.10
- [6] ShadeR5 3Dモデリングマスター 2001年(株)ソーテック社

- [7]ShadeR5 3D レンダリングマスター 2002 年
(株)ソーテック社
- [8]次世代メディアクリエイター入門 2003 年
(株)カットシステム