

画像検索のための3Dインターフェースの検討と課題

小池真由美*1 伊藤学*2 池田佳代*3 日高宗一郎*4 青木輝勝*2

*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。検索に際するインデキシングコストを抑えるとして、本稿で対象とする検索方式は、検索者が検索したい画像を手元に持たず、自らが描画した画像を元に画像データベースを検索する場合である。検索精度を高めるためには、この入力インターフェースに課題があると仮定し、3Dの機能を利用した3Dインターフェースを適用した場合の検索の有効性を示す。

A study on 3D Interface for Image/Video Retrieval and its Issues

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

*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. This paper focuses on the retrieval method in which the retriever does not have any target picture at hand, and tries retrieving such picture from picture database comparing with his/her own drawing. The authors assume that the subject to be solved is in input interface in order to achieve retrieval efficiency, apply 3D function to make 3D interface and show improvement in retrieval achievement compared to conventional interface.

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。現在実用化されている画像検索は、Google[1]に代表されるように、テキスト入力を中心だが、画像にテキストが付与されているという前提なくしては成り立たない。その膨大な作業を想像すれば、画像内容から直接、検索したい画像を取得できれば、インデキシングコストが軽減でき、画像検索の効率に貢献することは容易に想像できる。

このような検索方法としては、目標画像を手元に用意し、それに類似した画像を画像データベースから検索する方式と、目標画像が手元にないが、目標画像に近い絵を検索者が描画し、画像データベースから検索する方式とがある。

本稿で対象とするのは後者の方式である。

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

描画方式の場合の課題は、ユーザが、検索したい画像のイメージを正確に描けない(表現できない)という点であるが、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る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。

「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[2]。

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/動画画像を検索する際に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。

このシステムの、3D入力インターフェースとしては、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等に入力させ、最終的な演算結果(検索結果画像)を得る。

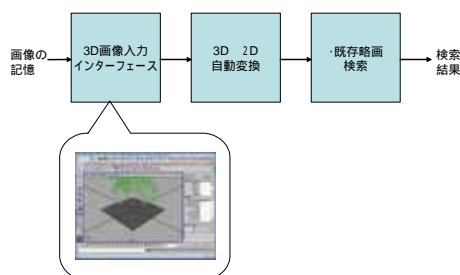


図2 3D入力インターフェース

2.3 実験範囲の検討

ところが、「3Dインターフェース」を既存のソフトウェアを対象に検討してみたところ、2Dインターフェースと同レベルの簡便さで入力できるものが極めて少ないということがわかった。

本稿で、目的としているのは、『あらゆる個人が使いやすい画像入力としての3Dインターフェース』であるため、画像描画を行う主体はあくまで画像描画の素人である。たとえ奥行き感の表現が画像上に実現され、検索したい画像により類似するため、3D描画ソフトウェアが必要だとしても、その習得に数日かかるのでは本末転倒である。

将来的には、複雑な要素を擁した自然画像そのものを描画できるような3D描画ソフトウェアを使用した検討も必要になる可能性はあるものの、ここでは、基本的な部分の検討、すなわち、3D入力インターフェースを使用することによって、2D入力インターフェースに比して検索効率上がるであろうと仮定される範囲から検討を始めた。

2.3 3Dインターフェースのための予備実験結果

この検討にあたり、過去の予備実験の結果を参照することとした。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による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、を筆者らは過去、[3][4][5]において検証し、以下の結果を導いた。

- (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個の物体の奥行き相関関係が把握しやすい。
- (3) 奥行きを把握しやすい3D配置において、白と黒の球の場合、白の球がより近く、黒の球がより遠く認識・表現される。ただし、これらの実験では、対象物数が2個となっており、2者の相関を前提とした、被験者の奥行き感を検証した形になっている。

そこで、新たな実験においては、図3のような3種の画像を使用して、対象物を1個に限った。そして、まず2D入力インターフェースでの描画実験を行った。

ここでは、同じ物体でありながら、異なる方向・角度からの撮影によって、すなわち、撮影されている面数が増加するにしたがって、描画時間が長くなること、描画が困難になることが実証された(平均時間: 1面 53秒、2面 122秒、3面 271秒)。特に、3面が撮影された画像に至っては目標画像の相似にもならない結果が殆どであった。[6]

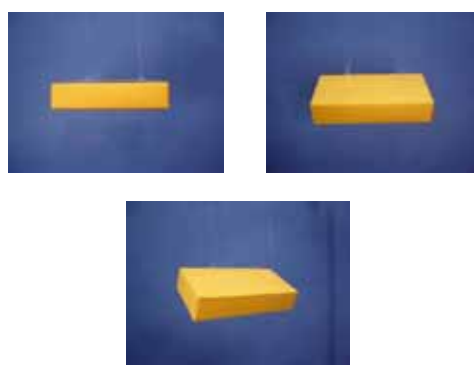


図3 2D入力実験対象画像

2.4 3Dインターフェース検証の前提と仮定

上記の実験結果を踏まえた場合、立体の概念を描画に落とし込むことの難しさが浮き上がってくる。ここを支援することのできるインターフェースがあれば、描画を用いた画像検索の精度向上に寄与できる可能性がある。

そこで、図4のような3D インターフェースを検討した。オレンジ色の部分が、既存の2D 描画による入力インターフェースと異なる部分になる。

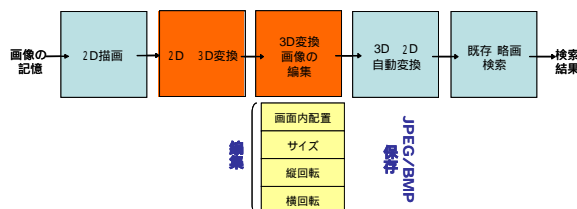


図4 3D インターフェース検証のプロセス

2D による描画を奥行きを持った3D 画像に変換する、または最初から奥行きをもった3D 描画を行い、その後、編集（回転や画面内配置、サイズ変更）によって、検索したい画像に近い描画にするのである。この編集部分で3D 描画の際に、より検索したい画像に近づくために有用なのは、奥行き部分のサイズ変更と、縦横の回転であろうと予測できる。

ここで、3D インターフェース検証を行うための前提と、2つの仮定を確認した。(図5)

前提として、ある画像を見たときに、

- ・ 3D 物体であること
- ・ どんな形状の3D 物体か（立方体か直方体か、など）

は、誰にでも判断がつく、ということ。

仮定1としては、

- ・ 2D で立体の描画を行うより、2D 描画を3D に変換した上で、物体を回転させるほうが、目標画像に近い『類似画像』を容易に作成できないか？

仮定2としては、2D 描画より3D 変換したもののほうが

- ・ 類似画像作成時間短縮
- ・ 検索精度向上

を見込めるのではないか？ ということである。

3. 3D インターフェース検証実験

3.1 実験準備

3.1.1 ソフトウェアの選定

前述の前提・仮定をもとに、2D 描画より奥行き感、立体感を表現できるであろう3D 描画が可能なソフトウェアを複数検討した。その結果、比較的一般に普及しており、パソコンでの描画経験が少ないユーザにも習熟が容易なツールとして、本稿では、Microsoft Powerpoint を選択した。

上記実験[6]でも、このソフトウェアを使用している。2D 描画が可能であり、ある特定描画ツール（三角形・四角形など）2D 描画から奥行きをつくることにより、3D 画像としての表現が可能である。

3.1.2 被験者の条件

また、実験に際して、被験者は、powerpoint を使用することができ、描画の経験もあるが、3D 画像表現を利用したことがないか、ほとんど無い者に限った。

そのため、Powerpoint の2D 描画方法および3D 描画方法について、基本機能操作5種（直線・矩形・矩形立体化・回転・奥行き変更）のが可能なように、事前に確認と簡単なトレーニングを行った。

3.2 実験概要

本稿で実施した実験の概要を示す。

- ・ 目的：直方体の3面が画像上に表現されている1枚の絵を、検索者としてのユーザが描画再現する上で、3D インターフェースが有効かどうかを確認する。すなわち、2D インターフェースと3D インターフェースの
 - (1) 描画時間
 - (2) 検索精度
 についての比較を行う。
- ・ 使用ソフトウェア：Microsoft PowerPoint
- ・ 実験使用画像枚数：1枚（図6）

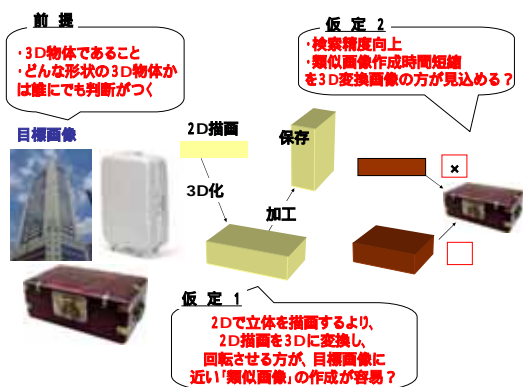


図5 3D インターフェース検証の前提と仮定

- (練習には別に1枚)
 ・被験者数：4名

3.3 実験手順

以下に実験手順を示す。

- (1) 被験者に右側に描画したい画像(以降“原画”と呼ぶ)を、左側にブランクを配置したページを見せる。(図6)

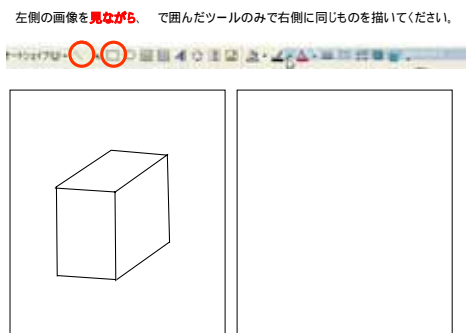


図6 手順(1)使用画像

- (2)被験者は、原画を見ながら左側ブランクに、同様の絵を2D機能のみを使用して描画する。
 (3)(2)と同様に、3D機能も使用して描画する。

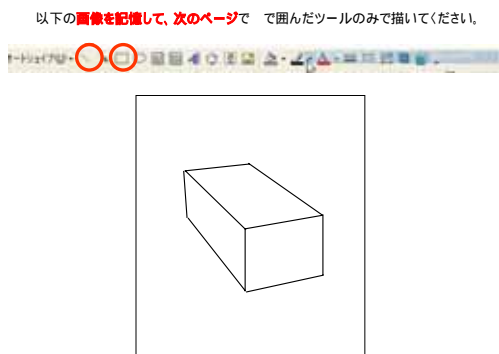


図7 手順(4)実験使用画像

- (4)(3)までは練習問題の位置づけである。本実験として、別の原画(図7)を用意し、原画を見た後、次ページに移り、原画を見ずに同様の絵を2D機能のみを使用して描画する。
 (5)(4)と同様に、3D機能も使用して描画する。
 上記実験において、被験者は、

- ・ 画面内における画像の位置
- ・ 画像の形状
- ・ 画像の大きさ

に留意する。尚、色は不要。
 また(2)~(5)のプロセスは各々時間を計測する。

3.4 評価方法

本稿では、原画と、描画像の類似度の基準を、対象画像の画面内の位置・形状・大きさに置くため、原画及び描画像を黒で塗りつぶし、

原画像、2D描画像、3D描画像それぞれについて画素毎に、RGBの和を算出し、各画素を、

- ・ 原画と2D描画像
- ・ 原画と3D描画像

の、差分をとり、画像全体の和が小さいものを、より類似しているとする評価を行った。

評価には、MATLAB version7.1(The Matworks Inc.)を使用した。

3.5 実験結果

3.5.1 描画時間比較結果

前述の実験手順(2)(3)に従い、原画を見ながらの描画の場合の時間を計測してみた。

ここでは、明らかに2D描画の方が時間がかかる傾向にある。(表1)

	2D	3D	差分
1	57	15	42
2	150	72	78
3	101	55	46
4	74	52	22
平均	95.7	82.0	

表1 練習問題の描画時間

次に、実験手順(4)(5)に従い、原画を記憶した状態からの描画の場合の時間を計測してみた。

ここでも、2D描画の方が時間がかかる傾向にある(表2)。

2D描画は業務でも使用するレベルの被験者のみを集めたが、3D描画は初めてであることを考慮すると、Powerpointを用いた3D描画程度であれば、習得はそれほど難航しないことがわかる。

(時間:秒)

	2D	3D	差分
1	42	30	12
2	105	54	51
3	45	70	-25
4	46	32	14
平均	68.3	83.5	

表2 本実験の描画時間

3.5.2 類似度比較結果

実験手順(4)(5)で実施した2Dインターフェースによる描画と3Dインターフェースによる描画を、前述の評価方法に従って、それぞれ原画と比較した結果を表3に示す(0 = 完全不一致、100 = 完全一致)。

(%)

	原画 - 2D	原画 - 3D	差分
1	88.8318	89.0461	-0.2143
2	92.1930	96.3347	-4.1417
3	90.1315	93.1053	-2.9738
4	95.5331	96.5150	-0.9819
平均	91.6724	93.7503	-2.0779

表3 類似度比較結果

この結果より、どの被験者も、2Dインターフェースによる描画より、3Dインターフェースによる描画の方が、原画像に類似した描画ができていて、ということが示された。

4. まとめ

本稿では、手描き画像を使った画像検索のためのインターフェースを模索する上で、3Dインターフェースを検討した。

検討には、描画時間および原画像との類似度の2点について従来の2Dインターフェースと比較するという手法をとった。

その結果、描画時間については3Dインターフェースに不慣れな被験者でも2Dインターフェースより短い時間で描画が可能であること、それにも拘らず、原画像との類似度については、3Dインターフェースがより高いという実証が得られた。

今後の課題として、本稿では直方体という限られた形状のみを対象とした実験であったため、他の形状においても3Dインターフェースが同様の

効果を発揮するのか、またどういった点で発揮するのかについて被験者も増加して検証を重ねたい。

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17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4人の被験者に心より感謝申し上げます。

文献

- [1] <http://www.google.com/>
- [2] 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
- [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] 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, 2005.3