



応
般

大量映像の分析と利用に基づく映像合成

岡部 誠

電気通信大学情報理工学部総合情報学科 / 科学技術振興機構さきがけ

撮影技術の進歩とインターネットの普及によって、映像はいまだかつてないほど身近になった。誰でもデジタルカメラで写真や動画を撮影できるし、共有サイト (Flickr, YouTube, ニコニコ動画など) にアップロードすれば世界中の人々と互いの作品を鑑賞し合える。このようなインターネット上の大量の映像データは映像鑑賞のみならず、映像製作サイドにもインパクトを与えている。実際、インターネット上の再利用可能な写真、動画は多くのホビーユーザーによって映像作品の製作に利用されている。テレビ局、CG プロダクションなどプロフェッショナルの現場でも独自に映像データベースを作って、効率よく検索し、加工・編集することで映像製作に役立っている。一方、我々 CG の研究者も映像データを大量に分析・利用することで、より使いやすい映像鑑賞ツール、映像製作環境の提案を行っている。以下、CG 分野で話題を呼んだ最近の画像データベースを用いた研究と、我々のグループが取り組んでいる動画の鑑賞、合成技術について紹介する。

画像検索技術に基づく画像合成

最初に Hays らが 2007 年に発表した「インター

ネットにある何百万枚もの写真を使って、自分の写真の気に入らない部分を自由に消そう」という研究について紹介する¹⁾。これは大量画像データの有効性をかなりはっきり示した研究で、CG 業界にも大きなインパクトを与えた。図-1 を見ていただきたい。(a) は綺麗な風景写真だが、手前に大きな家が写っている。ユーザは海が見たいのであり、こんな家は見たくない。そこで (b) のように、ペイントツールでこの家を塗り潰してしまう。するとコンピュータがこの消えた穴を埋めるのにちょうど良さそうな写真をインターネットから探してきて (c) のように表示してくれる。ユーザが適切な写真 (ここでは赤い四角で囲まれている) を選択すると、コンピュータはそれを使って穴を綺麗に埋めて (d) のような合成画像を作ってくれる。

この研究の肝は高速な画像検索である。(a) と似たような風景写真を高速で検索するために、データベース内の各画像はたった 4×4 ピクセルの GIST 特徴量 (元々の写真は 1000×1000 ピクセルだったりするので凄く圧縮されている。gist (ジスト) は主旨、要点の意) で表現されている。(c) でユーザが選んだ写真は、全自動で位置・スケールを合



図-1 (a) 元の写真 (b) 気に入らない家を消した (c) よく似た風景写真の検索結果一覧とユーザの選択 (赤い四角) (d) 合成結果

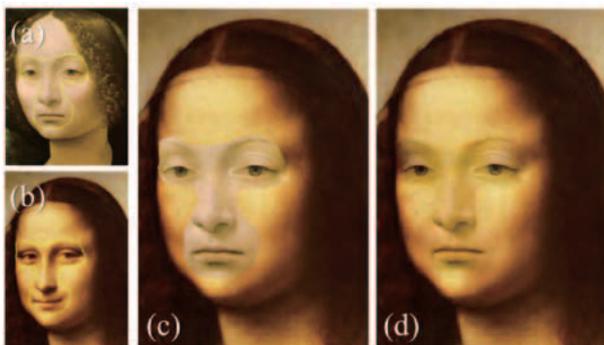
© 2007 ACM, Inc. Included here by permission.

わせつつ (b) 上に重ね合わされる。最後に境界部分がバレないように Poisson Blending を用いて合成される。Poisson Blending とは画像を切り貼り合成するための手法で、図-2 (a) の顔を切って (b) 上に貼っただけだと (c) のように境界がくっきり見えてしまうところを、(d) のように滑らかにつなげることを可能してくれる²⁾。手法自体は簡単で面白く、インターネット上で素晴らしい解説・プログラムも見つけられるので興味のある方は参照していただきたい。

さて、この研究のポイントであり、かつ、本稿で私が最もお伝えしたいメッセージは、この論文中に書かれた次のことだ。

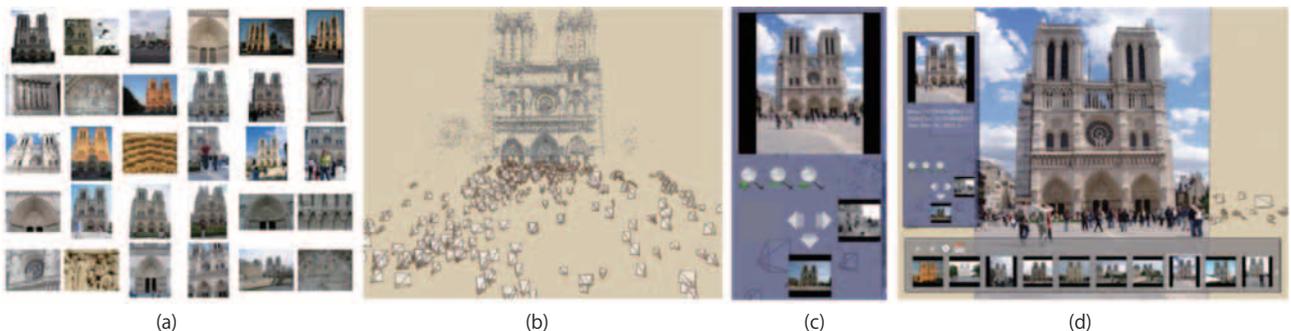
「同じ方法を 1 万枚の写真で実験していた頃は全然ダメだったけど、データベースに 200 万枚の写真を使うと上手く動き出した」

ここで使われている画像合成のアルゴリズムは上述のように、位置・スケール合わせと Poisson Blending という、どちらも比較的シンプルなもの



© 2003 ACM, Inc. Included here by permission.

図-2 Poisson Blending (a) の顔を切って (b) 上に貼ると見えてしまう (c) の境界を (d) のように消してくれる



© 2006 ACM, Inc. Included here by permission.

図-3 (a)「ノートルダム大聖堂」の検索画像一覧 (b) これらの写真を分析して得た 3 次元情報 (c) Photo Tourism のユーザインタフェース (d) Photo Tourism のシステム

だ。より良い結果を得るために、これらをもっと複雑なアルゴリズムにして……、というのが今までの研究開発のスタイルだった。だけど、この論文は「そこに心血を注いでもいいが、単純にデータをたくさん用意すれば、アルゴリズムはそのままでも結果が良くなったよ」と言っている。これが、大量の映像データに我々が注目する理由であり、データ駆動型アプローチの面白いところである。ちなみに、この論文の原題は「Scene Completion Using Millions of Photographs」であり、対象は scene、つまり風景写真のみである。このアイディアの風景以外への拡張も後で紹介する。

次に Snavely らが 2006 年に発表した Photo Tourism を紹介する³⁾。画像検索技術と 3 次元復元技術を組み合わせることで、あたかもその場所に行ってそこを移動するかのように、インターネット上の写真を鑑賞させてくれるツールである。図-3 を見ていただきたい。今、ノートルダム大聖堂の写真を見たい。普通に Flickr や Google の画像検索を使って「ノートルダム大聖堂」と言葉で検索すれば (a) のような一覧が表示される。これで十分と言えば十分なのだけど、たとえば「もっと近くで撮った写真が見たい」「寺院を右側から撮ったような写真はないかな？」などと思った場合、この一覧を 1 つ 1 つ注意深く見て探さないといけない。そこで Photo Tourism では各写真がどの位置から撮影されたものを 3 次元的に分析する。(b) がその分析結果で、各写真を撮影したカメラが小さな四角錐で地面に表示され、一番奥に 3 次元化されたノートルダ

ム大聖堂が見える。この3次元情報を利用することで、(c)の矢印ボタンのように「このノートルダム大聖堂の正面写真から、もう少し右に移動して撮った写真をください」や、(d)の下段の写真一覧のように「この写真の周辺で撮影された別の写真を全部ください」といった鑑賞ツールを提供できる……、と文章で読むよりビデオやデモを見ると一目瞭然なので、ぜひ「photo tourism」でGoogle検索していただきたい。あたかもノートルダム大聖堂の周辺を3次元的に移動しているかのように写真が鑑賞できる。

ここで肝となる技術は structure-from-motion (SfM) である。人間もカメラ2つでステレオ視することで3次元的奥行きを感じているが、このカメラの数は多ければ多いほど復元できる3次元情報も詳細になる、というのがこの技術である。同じ対象（ここではノートルダム大聖堂）を異なる位置・方向から撮影した写真をたくさん用意してSfMを適用することで、(b)に示すように撮影対象自体を3次元化し、同時に各写真を撮影したカメラの位置・方向も3次元的に知ることができる。

次に紹介するのは、Leeらが2011年に発表したお絵描き支援手法 ShadowDraw である⁴⁾。この手法は、画像をリアルタイムに検索してお手本として表示することで、お絵描きを支援する。図-4(a)がシステムを使って絵を描いているところである。ここでユーザが描いたのは濃い線だけで、薄い線（長髪の人顔に見える）はユーザの描いた線をクエリにして検索した画像を、システムが勝手に重ね合わ

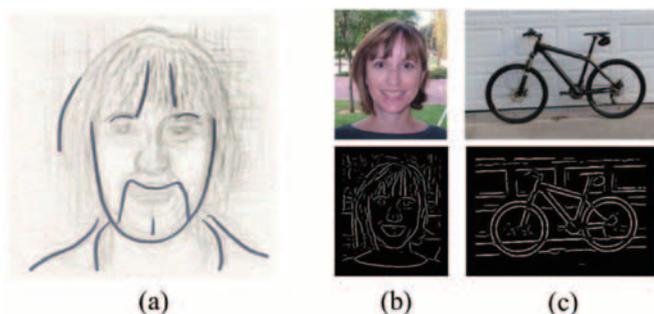
せて表示しているものである。ユーザはこの薄い絵をなぞるように絵を描いていくことができる。これもYouTubeで「shadowdraw」と検索するとビデオが出てくるので、見ていただくと一目瞭然である。

ShadowDrawは、インターネットから画像を収集し、各画像にフィルタをかけて抽出したエッジ画像をデータベースとして持っている（図-4(b), (c)）。技術的な肝はユーザが今描いている絵（濃い線）に基づくデータベースの高速な検索である。ここでは、各画像を将棋盤のように9×9の領域に区切って、各領域でBiCEと呼ばれる特徴量を記述している。BiCE特徴量はエッジ画像の大事な情報をなるべく失わず、かつ、小さなデータ量に圧縮して表現することに長けているので、高速検索を可能にしてくれる。検索されたエッジ画像はユーザの絵との類似性を考慮しつつ、似たものは濃く、似ていないものは薄く重ね合わせて、図-4(a)のように表示される。

論文では実際に何人かに使ってもらって「システムを使うと本当に絵が上達するのか？」を評価していて面白い。実験には絵が上手い人から下手な人までさまざまな人が参加しているが、「上手くはないが下手でもない、まあまあ絵心のある人を支援することができました」という結果が書かれている。つまり、もともと絵が上手い人は支援の必要がなかったし、このシステムで支援しても効果がないほど下手な人もいた、ということである。

人・物体の認識に基づく画像合成

写真中の人や物体を抽出して画像編集に使う技術として、Lalondeらが2007年に発表したPhoto Clip Artを紹介する⁵⁾。図-5(a)がシステム画面である。背景画像を読み込み、メニューから人や車など、配置したい物体のジャンルを指定する。今、ユーザは横断歩道に人を置きたいので、システムはこのシーンに置くのに良さそうな人をデータベースから探してきて、右側に一覧表示している。ユーザはこの中から適当なものを選んで配置することで、効率よくシーンをデザインできる。



© 2011 ACM, Inc. Included here by permission.

図-4 (a) ユーザの描いた絵（太い線）とシステムが提示するお手本（薄い線） (b) 顔の画像とそのエッジ画像 (c) 自転車の画像とそのエッジ画像

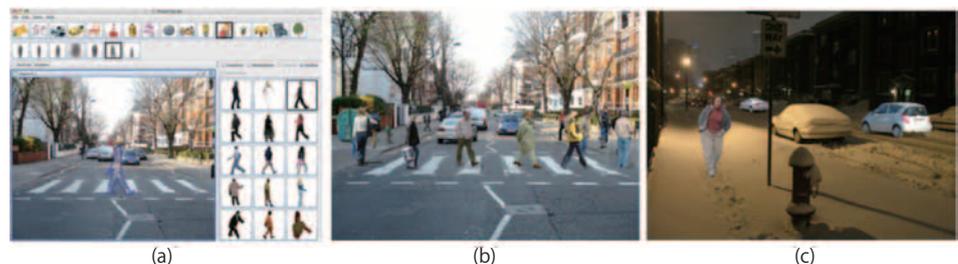
この研究の肝は人や車のサイズ、また、照明条件を分析して、このシーンによく合う人や物体を出してくれるところだ。たとえば、図-5(a)は日陰になった街路の写真だが、同様の照明条件の写真のみから人々を抽出している。実際に、人や車を配置したのが(b)で、自然なシーンができている(何も知らずに見ると合成だと気付かない……かもしれない)。一方、何も意識せずに適当にシーンに人を配置すると不自然な結果になりやすい。たとえば(c)のように夜の背景に昼間の写真から切り取った人や車を入れると、照明条件が違うので明らかに不自然な結果ができてしまう。この昼間の人や車を夜の人や車に見せるべく、画像を無理やり編集することも不可能ではない。実際、写真の中の照明を編集する研究は数多くある。しかし、インターネット上に大量の写真があって、しかも、今作りたいシーンは誰でもいいから人が歩いていてくれれば十分なのだから、照明編集のような難しい問題を解くよりも、適切な画像を検索してきた方が話が早い、というのがこの研究の哲学である。

余談だが、人や物体を切り抜いてデータベースにまとめる作業は「Label Me」に基づいている。Label Meはインターネット上のアプリケーションで、写真を切り抜いて、各パーツに「人」「車」などのラベル付けを手作業で行えるようにしたものだ。切り抜きやラベル付けは現在も盛んに研究されている分野だが、全自動で完璧な結果を出すのはいまだに難しい。そこで簡単なアプリを作って公開しておき、世界中の暇人に頼んで作業してもらった方が早い、というのがLabel Meの思想である。大量の写真、大量の人に仕事を依頼でき

るインターネットの出現によって、難しかったコンピュータビジョンの問題もシンプルに解けるといふ一例である(「label me」で検索すると遊べます)。

次にChenらが2009年に発表したSketch2Photoという研究を紹介する⁶⁾。目的は先ほどと同様にインターネット上の写真から人や物体を切り出して自分の好きなシーンを作ることだ。図-6(a)では入力として言葉で欲しい物体を指定しつつ、スケッチでシーンの雰囲気を描く。この例では背景に「sunset beach」、スケッチで「sailboat」の形や「wedding kiss」をするカップルのポーズを指定している。この入力を元にコンピュータは自動で画像検索、切り出し、合成までを行い、図-6(b),(c),(d)に示すような、いくつかの候補を作ってくれる。

この技術の肝は、現行のCG技術の限界を理解した上で、合成してうまく見えるような画像の組合せをインターネットから探してくるところにある。図-7に示す例は、システムが上段の写真からヒツジを切り取って、下段で別の草原の上に合成している。(a)や(b)は良い結果で、パッと見では合成だと気付かないかもしれない。(c)はちょっとまずい。合成の影響でヒツジが黄色くなってしまっている。(d)はもっとまずい。ヒツジの切り取りに失敗し、



© 2007 ACM, Inc. Included here by permission.

図-5 (a)Photo Clip Artのシステム画面 (b)写真内に人や車などを配置して作ったシーン (c)照明条件の合わない人や車を合成してしまった失敗例



© 2009 ACM, Inc. Included here by permission.

図-6 (a)言葉とスケッチによるユーザ入力 (b)(c)(d)システムが作ったシーン

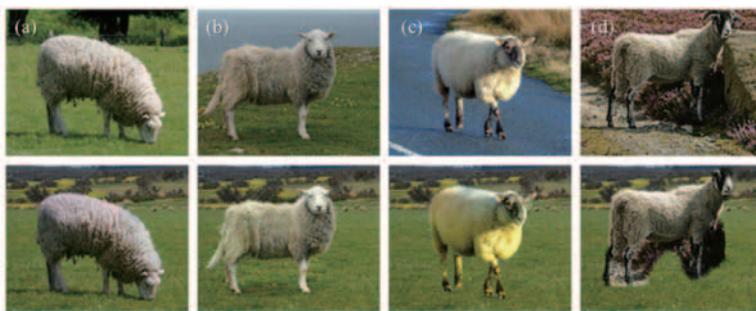
足元の紫の花、白い砂利道が残ってしまっている。

さて、これらの結果はヒツジを切り取って、Poisson Blending で合成したものだが、難しいのはヒツジの切り取りである。実は、この4つのすべての例において、システムはヒツジを綺麗に切り取ることは失敗している。(d) ははっきり分かるが、上手くいっているように見える(a)や(b)でも、特にヒツジの足の間をよく見ると元画像の一部が残っている。ただ(a)や(b)のヒツジはもともと緑の草原にいたので、合成先の緑色の草原と相性が良かったのだ。つまり、切り取りミスで多少ヒツジのフチに緑色が残っていたとしても、Poisson Blending (図-2 参照) が綺麗につないでくれたのである。一方(c)はもともと青い道路にいたヒツジなので、切り取りミスで残った青をムリに緑に変換しようとした結果、ヒツジが変色してしまった。(d)では切り取りミスで残った花や砂利道が草原と色的に違い過ぎるため、システムはPoisson Blending を適用しないという判断をしている。

(c)や(d)のヒツジも、もっと綺麗に切り取ることで上手く合成できる。しかし、画像の切り取りを全自動で完璧に行うのはいまだ難しい問題である。た

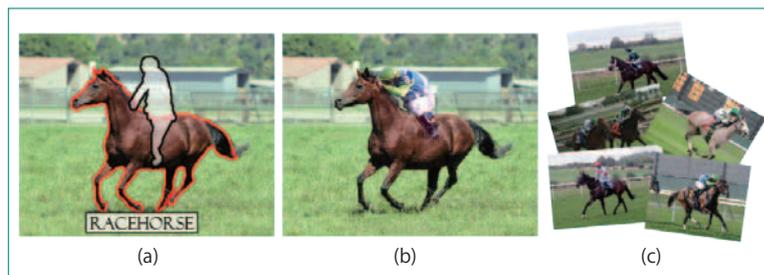
だ、(c)や(d)のヒツジにこだわるのであれば、もっと合成しやすい(a)や(b)のヒツジを見つけてくれば良い。画像合成技術1つ1つの改良は難しいが、現行の技術でも上手くいくヒツジと背景のペアを、大量の画像の中から探してくるのはもっと簡単、というのがこの手法のアイディアである。

さらなる発展としてGoldbergらが2012年に発表したData-Driven Object Manipulation in Imagesでは、検索して切り取った人や物体を多少変形させることで、さらに柔軟な合成処理を可能にしている⁷⁾。図-8 (a) では馬の上にジョッキーを乗せるため、形をスケッチして「Race Horse」という言葉を与える……あたりまではSketch2Photoと変わらない。(c)にインターネット上でシステムが見つけた、ちょうど良さそうなジョッキーの乗った馬の写真がある。ところが、これらのジョッキーを切り取ってそのまま貼り付けただけでは、この馬の上にピッタリと収まらない。そこで、システムは多少の変形を加えてジョッキーの姿勢を曲げて修正することで、馬にピッタリ乗ったジョッキーの画像(b)を合成している。



© 2009 ACM, Inc. Included here by permission.

図-7 ヒツジを切り取って別の背景の上に重ね合わせる例



© Chen Goldberg

図-8 (a) 馬にジョッキーを乗せるためのスケッチと言葉のクエリ (b) 変形されて馬にピッタリ乗ったジョッキー (c) ジョッキーの候補画像

画像・動画データベースに基づく我々の取り組み

ここからは、我々のグループ（電気通信大学 尾内・岡部研究室）が行っている画像・動画データベースに関する研究を紹介する。最初に川手らが2012年に発表した「動画鑑賞のためのスケッチインタフェース」に関する研究を紹介する。今、図-9 (a)のように、車が左から右へ走るシーンを見ているとする。このとき「この車が手前へ走ってくる、(b)みたいなシーンがあったなあ、あれを見たいなあ」と思うことがある。普通だと、早送り、巻戻しボタンや再生バーを操作してシーンを探す。それでも見つからない場合はYouTubeへ行行って検索するかもしれない。しか

し、ここではもっと迅速に目的のシーンへ移動する手法を提案する。

ユーザの見た (b) のシーンとはつまり、今見ている (a) の車を、クルッと回して前を向かせたようなシーンのことである。そこで、このクルッと回す感じを (c) のように緑の矢印でもって、ユーザにスケッチしてもらおう。「回転して、前後のウィングがちょうどこの辺に来ると、私の見たいシーンになるなあ」という感じである。すると、システムが即座に目的のシーンを検索してくれる。デモ動画は「川手 岡部 尾内」で検索するとトップに川手氏の Web サイトが出てくるので、そちらでご覧いただきたい。

この技術の肝は、画像検索と軌跡検索を組み合わせたところにある。画像検索だけだと、(a) の画像が与えられても (b) の車を直接検索することは難しい。この様子を図-10 に示す。画像検索でよく用いられる、局所的な画像特徴のマッチングを試みた例である。(a) に示す車は左右反転しているものの、車自体は似たような姿勢をしていてマッチングがとれやすくなるので検索することができる。一方、今ユーザが見たいと思っている (b) は、局所的なマッチングは3つしかとれなかったため通常は検索結果とならない。

しかし、この (a) と (b) の関係に注目する。実は (b) の車が手前に走ってきて左側へカーブしていくのが (a)、つまり (a) と (b) は一続きのシーンなのだ。そこで我々のシステムは、まず局所マッチング

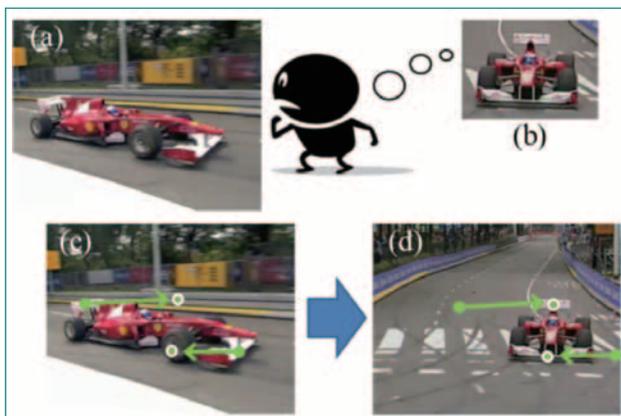


図-9 (a) 左から右へ走る車を鑑賞中に (b) のような手前に走ってくる車のシーンを見なくなった (c) 前後のウィングをつまんでクルッと回すイメージで緑の矢印を描くと (d) の目的のシーンが検索される

に基づいて (a) を検索し、次に (a) を起点に早送り、巻戻しを自動で行いながら「ユーザがスケッチした緑の矢印みたいな軌跡を描いて前後のウィングが動いたようなシーンはないかしら？」とチェックする。すると、(a) を少し巻戻したところに (b) が見つかるというわけである。

次に岡部らが 2011 年に発表した「ビデオデータベースを用いて流体画像をアニメーションさせる手法」を紹介する。図-11 (a) はキャンプファイアの写真、(b) は 400 年くらい前に描かれた油絵である。1 枚の絵であれば、(b) の油絵のように、アーティストは独自のタッチで詳細に描き込むことができる。しかし、そのようなシーンのアニメーションを作ろうとすると、流体シミュレーションを走らせて 3 次元 CG で再現するのも難しいし、かといって何枚も同じような絵を描いてパラパラアニメ化するのもとても大変である。そこで、1 枚の絵の持つ雰囲気、質感をなるべく損なわないようにアニメーションを作るため、我々はビデオ検索技術と画像合成技術を組み合わせる技術を開発した。

システムは図-12 (a) に示すように、流体ビデオのデータベースを持っている。ユーザが指定するの

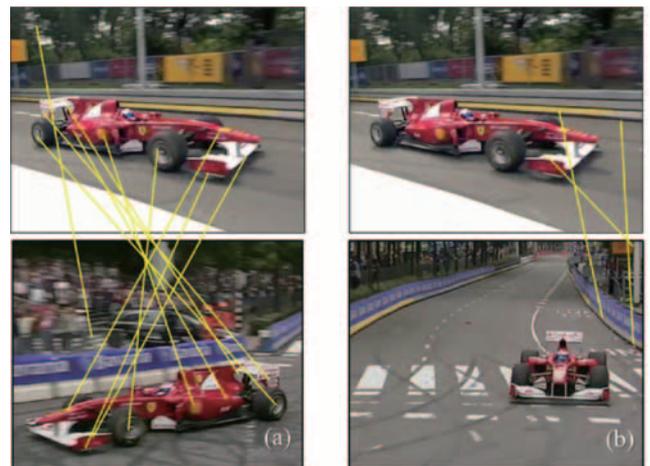


図-10 画像検索で用いられる局所的な画像特徴のマッチング



図-11 (a) キャンプファイアの写真 (b) 川の急流を描いた油絵

は3つで、(b)の流体画像とその上の大体の流れ方向を指定するオレンジの矢印、(c)の流体領域である。後は自動的にシステムが(d)の流体アニメーションを作る。「岡部 ビデオデータベース」で検索するとデモ動画が見られるのでご覧いただきたい。

この技術の肝は、局所的な動画検索である。

図-13は、3つのビデオデータを用いて小川の写真をアニメーションさせようとしている。この3つの小川のビデオは写真の小川とはまったく異なるものだ。しかし、局所的に見ると似ている部分もある。たとえば、黄色、オレンジ、青の丸印で囲んだ領域では、それぞれビデオと写真の間で水しぶきの調子が似ている。このような類似性に基づいて、各丸で囲んだ部分のビデオを写真上へコピーする、というのがアイデアである。写真上のすべての場所へ、いろんなビデオから部分的なコピーをしつつ、画像合成技術を使って切れ目が見えないように合成することで最終的なアニメーションを得ることができる。

次に紹介するのは音楽に関する研究で、山本らが2012年に発表した「入力音楽をあたかもバンド演奏しているかのように見える動画の生成技術」である。入力音楽に合わせて、その音楽の演奏動画を半自動的に生成し、演奏を耳と目の両方で楽しむことができる。この手法ではピアノ、ギター、ベース、ドラムなどの演奏動画を収集して動画データベースを作っている。図-14のDatabaseに示すように、

動画のオーディオトラックをあらかじめ分析することで、図中の(b)などの音符の発音タイミングを抽出しておく。この発音タイミングは実際に演奏者が演奏しているであろう楽譜上の音符の位置に相当するようなものだ。今、図-14のInput Musicに示すような入力音楽が与えられたとすると、システムは同様のアルゴリズムで図中の(a)に示すように発音タイミングを抽出する。この発音タイミングを特徴量として動画データベースを検索することで、類似した発音タイミングを持つ動画を検索して図-14中の「???」部分に割り当てる。ここでは(a)と(b)は類似しているので、(b)の動画を割り当てれば、(a)を演奏しているように見えるに違いない、というわけである。この作業を入力音楽のすべてのパートに繰り返すことで、演奏動画を生成する。デモ動画は「yamamoto okabe onai」などと検索するとトップに山本氏のWebサイトが出てくるので、そちらでご覧いただきたい。

図-15に合成動画のフレームをいくつか示す。音楽をただ聴くだけだと、それほど注意して聴いていなかった低いベースの音も、映像中のベース奏者の指の動きと一緒に聴くとより鮮明に聴こえてくる。映像（偽モノですが）が音楽の楽しみ方に与える影響について、引き続き分析と実験を行っている。

最後に紹介するのは、水井らが2012年に発表した「アマチュア演技者の動作を格好良く見せるよう

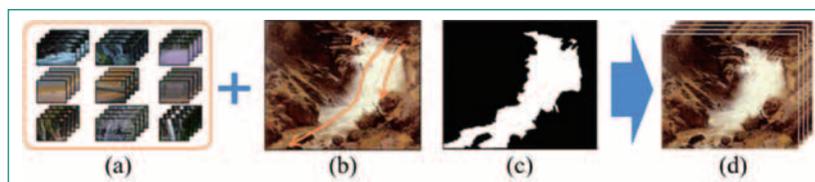


図-12 システム概要

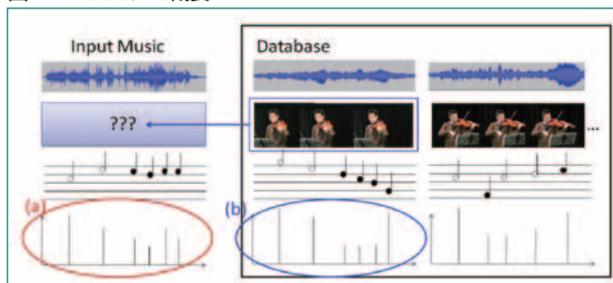
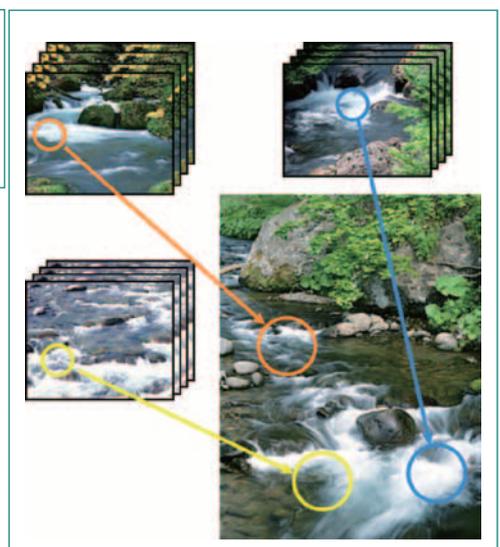


図-14 入力音楽とデータベース。両者において、音符の発音タイミングを自動的に抽出して検索を行う



▶ 図-13 ビデオの部分的なコピーと画像合成によって小川のアニメーションを作る

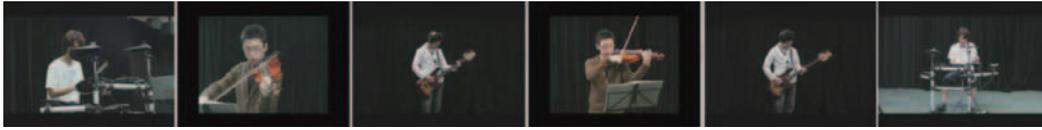


図-15 ここでは、ビートルズの「レットイットビー」を入力音楽としてバンド演奏を合成している



図-16 プロの動作を抽出してアマチュアの動画に適用することでアマチュアを格好良くする

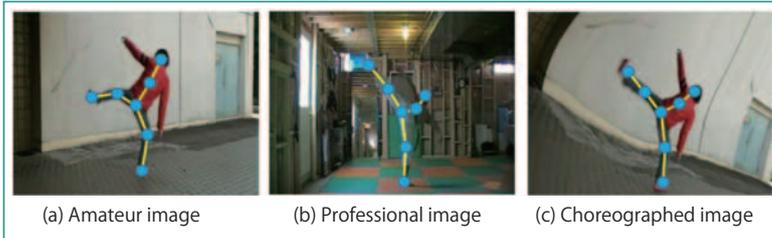


図-17 スケルトンを用いた動作の追跡と変形への適用

に動画を編集する手法」である。プロの演技者はダンサーでもアクションスターでも非常に格好良い。体は柔らかく筋力があり動きにキレがある。ところが、我々がそれに憧れて真似ようとしても上手くいかない。体は硬く筋力はなく動きにキレがないのである。そこで、コンピュータに手伝ってもらい、せめて動画の中でだけでも格好良くなろう、という願いに基づいた研究である。

図-16 (a) では、アマチュアが回し蹴りに挑戦している。しかし、(b) のプロのように足は高く上がらないし、動きも遅くて格好悪い。そこで、このプロの動作を抽出してアマチュアの動画に適用することで編集したものが(c)である。これで、足は高く上がっているし、動きも速く格好良くなった。「mizui okabe onai」で検索するとトップに水井氏の Web サイトがあるので、ぜひデモ動画をチェックしていただきたい。

技術的には、アマチュア動画とプロ動画の両方の動作を分析して抽出(図-17 (a)と(b))して、前者を後者に合わせて編集する(図-17 (c))。動作の分析は図のように、人物の動きをスケルトンで追跡することで行う。この際、全自動でコンピュータが動きを完璧に追跡することは難しいので、間違いが起

こった際はユーザがマウスでスケルトンを編集できるようなツールを作っている。スケルトンが抽出できたら、(a)のスケルトンが(b)のスケルトンに重なるように(a)の画像を変形する。すると(c)ができ、足が高く上がったように見える(背景もそれに合わせて歪んでいるが、人物のみを自動で切り出して最終結果を合成する)。

参考文献

- 1) Hays, J. and Efros, A. : Scene Completion Using Millions of Photographs, In Proc. of SIGGRAPH 2007.
- 2) Pérez, P., Gangnet, M. and Blake, A. : Poisson Image Editing, In Proc. of SIGGRAPH 2003.
- 3) Snavely, N., Seitz, S. M. and Szeliski, R. : Photo Tourism : Exploring Photo Collections in 3D, In Proc. of SIGGRAPH 2006.
- 4) Lee, Y. J., Zitnick, C. L. and Cohen, M. F. : ShadowDraw : Real-Time User Guidance for Freehand Drawing, In Proc. of SIGGRAPH 2011.
- 5) Lalonde, J. F., Hoiem, D., Efros, A. A., Rother, C., Winn, J. and Criminisi, A. : Photo Clip Art, In Proc. of SIGGRAPH 2007.
- 6) Chen, T., Cheng, M. M., Tan, P., Shamir, A. and Hu, S. M. : Sketch2Photo : Internet Image Montage, In Proc. of SIGGRAPH Asia 2009.
- 7) Goldberg, C., Chen, T., Zhang, F. L., Shamir, A. and Hu, S. M. : Data-Driven Object Manipulation in Images, In Proc. of Eurographics 2012.

(2012年2月20日受付)

岡部 誠 (正会員) m.o@acm.org

2008年に東京大学で博士(情報理工学)を取得。マックスプランク研究所を経て、2010年から電気通信大学助教/科学技術振興機構さきがけ研究員。大量の映像データに着目しつつ新しい映像鑑賞ツールや映像製作環境の研究開発に従事している。