社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

［パネル討論］データ工学とメディア理解との融合 ～大量データにおけるパターン情報を扱う新しい枠組みは？～

井手　一郎[†,††]　　新田　直子[†††]　　尾関　基行[††††]　　中村　裕一[††††]

† 名古屋大学大学院情報科学研究科　〒 464-8603 名古屋市千種区不老町
†† 国立情報学研究所　〒 101-8430 東京都千代田区一ツ橋 2–1–2
††† 大阪大学大学院工学研究科　〒 565-0871 大阪府吹田市山田丘 2-1
†††† 京都大学学術情報メディアセンター　〒 606-8501 京都市左京区吉田本町
E-mail: †ide@is.nagoya-u.ac.jp, ††ide@nii.ac.jp, †††naoko@comm.eng.osaka-u.ac.jp,
††††{ozeki,yuichi}@media.kyoto-u.ac.jp

あらまし　従来，パターン認識・メディア理解の分野では，映像，ウェブ，ブログなどの時間軸を持った大量のデータを対象とした構造化・検索・要約・製作に関する研究が行われてきました．一方，データ工学の分野では，大量データ・時系列データの構造化・データパターンの発見などに関する研究が盛んに行われています．しかし今後，映像，ウェブ，ブログを対象とした構造化・データパターンの発見による検索・コンテンツ製作といった両分野を融合した研究分野の発展のためには，パターン認識・メディア理解側の考える「検索」とは何か，逆にデータ工学側の考える「パターン」とは何か，といった融合の土俵となる概念の定義について互いに明確にし，共通認識を図ることが重要となると考えられます．そこで,「大量データ」「時系列データ」をキーワードに，その構造化・検索・要約・製作を実現するための，データ間の関係性・規則性といったパターン抽出，収集・蓄積，可視化の方法などについて，現在研究されている方々をパネリストとして迎え，議論します．

キーワード　大量データ，時系列データ，パターン，検索

[Panel Discussion] Integration of data engineering and pattern understanding - How shall we handle patterns in massive data? -

Ichiro IDE[†,††], , Naoko NITTA[†††], Motoyuki OZEKI[††††], and Yuichi NAKAMURA[††††]

† Graduate School of Information Science, Nagoya University, 464–8603, Japan
†† National Institute of Informatics, 101–8430, Japan
††† Graduate School of Engineering, Osaka University, 565–0871, Japan
†††† Academic Center for Computing and Media Studies, Kyoto University, 606–8501 Japan
E-mail: †ide@is.nagoya-u.ac.jp, ††ide@nii.ac.jp, †††naoko@comm.eng.osaka-u.ac.jp,
††††{ozeki,yuichi}@media.kyoto-u.ac.jp

Abstract　Structuring, retrieval, summarization, and editing of massive chronological data, such as video streams, web pages, and blogs, have been active research topics in the pattern recognition field. On the other hand, in the data engineering field, structuring and pattern discovery targeting massive chronological data are the current hot topics. In order to support the development of an integrated research field of these two fields such as retrieval and content editing based on the data structuring and pattern discovery, it is important to establish a clear consensus between both fields on the definitions of common issues, that is, “What is a ‘pattern’ from the data engineering perspective?” and “What is expected for ‘retrieval’ from the pattern recognition perspective?”. In this panel, cutting-edge researchers in both fields will discuss issues on the methodologies to extract patterns such as relations and regularities in data, to collect, archive, and visualize ‘massive’ and ‘chronological’ data, which are necessary to realize their structuring, retrieval, summarization, and editing.

Key words　Massive data, chronological data, patterns, retrieval.

■ はじめに

今月の PRMU と DE の共催研究会では,「データ工学とメディア理解との融合 〜大量データにおけるパターン情報を扱う新しい枠組みは？〜」と題して，パネルディスカッションを企画した．DE 側，PRMU 側双方から新進気鋭の研究者を集め，各々の立場から，取組んでいる研究課題，これから期待する技術などについて意見を交すことにより，今後の研究や DE の研究者と PRMU の研究者の協力の可能性を探る予定である．以下では，PRMU の分野で行われてきた映像処理を題材にしてその紹介をするが，DE との関連はさらに多くの面に及ぶことに注意されたい．

□ パネル討論の背景（ＰＲＭＵ 側から）

映像データの処理は主に 1990 年代から行われてきている．それまでにも動画像処理は盛んに研究されてきたが，動画像処理とは目的が異なり，写っている対象の幾何的な計測や認識よりも，映像データ自身の時間的，意味的構造の解析に焦点があてられているのが特徴である．このような問題意識から，当初は，放送映像を録画したデータからショット替わり（シーンチェンジ）点を抽出し，ショット毎への分割，ショットを基にした構造解析，一覧提示，類似検索などが盛んに行われてきた．徐々に，一覧や検索以外の種々の応用に研究対象が広がってきた．

3 年前（2003 年 7 月）に「映像・マルチメディアとパターン認識・理解」というテーマセッションを開催した際には，将来の方向性も含め，図 1 にあげたような，要約，ポータル，大量データ，リアルタイム性，インタラクション，撮影,（コンテンツ）製作支援，個人適応などがキーワードとしてあがっている．各々の研究例（参考文献など）はその予稿 [1] に譲るが，大まかに眺めると，より広範囲の研究が行われるようになってきたことがわかる．

□ データ工学とパターン認識・メディア理解の接点

現在の状況は 3 年前と大きくは変わっていないと思われる．しかし，計算機周辺技術が進むにつれて，より一般的，より大量のデータが扱われるようになってきた．放送映像が既に数年に渡って蓄積されていたり，ライフログで代表されるような長時間の撮り流し映像や，多数のカメラで同時に撮影したような映像が題材とされたり，また，WWW が大量の画像・映像のリソースとして注目されている．このように，データ工学とパターン認識・理解技術の接点に踏み込む研究がますます多くなっている．

このような問題意識から，PRMU 側で集まり，このパネルセッション企画を考えた結果，DE の研究者が考える「パターンとは何か」，PRMU の研究者が考える「検索とは何か」を切口に，お互いの共通点，差違を考えることから始め，基礎理論，技術，応用について議論することにした．

映像データの蓄積，検索，ブラウジングなどはデータ工学で扱われてきた問題にそのまま当てはまる．映像中の類似シーンの検索は言うまでもなく，映像の要約一覧（ブラウジング）は，データの可視化やナビゲーションに，映像データの構造化やパターン学習はデータマイニングの手法に深く関わる．構造やセマンティクスの扱い方は，両者に共通する深い問題である．このように，両方の側の抱えている課題に接点は多く，良いツールを共有することの意義は大きい．

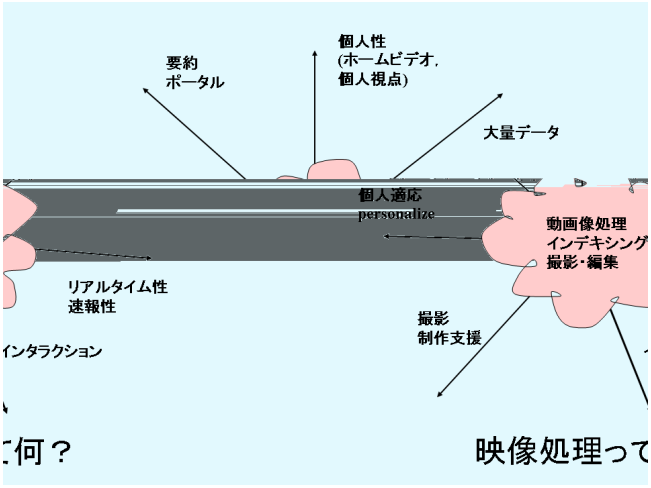

図 1　映像処理の広がり

さらに，ただ単に共通点を探るだけではなく，差違も興味深い．映像処理における早見（スキミング）の考え方はデータ工学ではどのような概念に対応するのか，ユーザの嗜好に対する個人適応はどうかなど，比較することにより興味深い知見やアイディアが得られるかもしれない．また，PRMU の分野では，画像・映像データの撮影やコンテンツ作成を行う研究が行われているが，撮影時に画像認識技術や他のセンシング技術を使って，付加情報の多いデータにすることも重要な課題となっている．大量のデータがまずあるところから出発するのではなく，「どのようなデータをつくれば良いのか」という点での議論も興味深い．

これらに限らず，事例ベースによるパターン認識など，多くの面白い課題があり，活発な議論となることを期待している．

□ パネリスト

今回のパネル討論では比較的若い研究者に声をかけ，これらのテーマについて意見や夢を語って頂くことにした．DE 側からは有村（北大），片山（NII），櫻井（NTT），豊田（東大），PRMU 側からは，井手（名大），尾関（京大），新田（阪大），中村（京大）を予定している．フロアからも活発なご意見や議論への参加をお願いしたい．

中村裕一 （京都大学）

■ 半構造データのための高速なマイニング技術（有村:DE）

□ スタンスと現状

データマイニング分野において，時系列やグラフデータのような構造をもつデータから，データを特徴付ける離散的なパターンを発見する高速なアルゴリズムの研究に取り組んでいます．特に，2000 年頃から，半構造データマイニング (semi-structured data mining) という名前で，それまでに個別の分野で研究され

り上げるよりも，事例ベースやテンプレートベースに解析・検証する方が効率的かつ高精度であったり，網羅的に学習セットを用意できるようになったことから解析手法が従来よりも汎用的かつ高精度になっているようです．大規模な映像コーパスが利用できるようになることで，映像を対象とする技術においても同様のパラダイムシフトが起きる可能性があると考えられます．そのような変革において，データベースシステム技術が，処理の効率化（特に，スケーラビリティとオンライン処理）の面で有意な貢献ができるのではないかと期待しています．

■ 多次元ベクトルと時系列データのための高速探索技術（櫻井:DE）

□ スタンスと現状

データベース分野において，多次元ベクトルの近傍探索や時系列データのマッチングの高速化について取り組んでいます．これらは映像のみに焦点を絞った研究ではありませんが，映像も時系列データの一つと見なせますし，画像処理によって得られた特徴が多次元ベクトルとして表現される場合もありますので，関連しているのではないかと思います．

多次元ベクトルの高速近傍探索

問合せとして入力されたベクトルデータと類似したデータを，大量の多次元ベクトルのデータベースから，高速に検索するための技術について取り組んできました [9]．

また，距離尺度としてユークリッド距離の他に，マハラノビス距離であっても高速に近傍探索ができるような技術にも取り組んでいます．検索システムの利用者が問合せ毎に異なる重み（異なる行列）を指定しても，その指定されたマハラノビス距離で高速に類似検索できるような技術です．

時系列データの高速マッチング

最近，データベースやデータマイニングの分野では，データストリーム処理という技術があります．これは，ネットワークから流れてくる大量の時系列データ（データストリーム）をディスクに格納せずに，処理するような技術です．

データストリーム処理の中でも，遅延相関検出について取り組んでいます [10]．これは，複数のストリームがネットワーク上を流れているときに，どのストリームとどのストリームが相関しているのかを検出し，さらに，それら相関があるストリームの各ペアにはどれくらいの遅延があるのかについても検出するような技術です．

その他に，動的計画法による時系列のマッチング（DP マッチング）や HMM （隠れマルコフモデル） について，精度を犠牲にすることなく，高速にパターン検出する技術についても取り組んでいます．

□ パターン認識／抽出技術に期待するもの

いくつかの例外はあるものの，パターン認識の分野では主として認識精度を高めることに重きをおかれてきました．データベースの分野では，比較的シンプルなパターン認識技術に着目し，その高速化に取り組んでいます．パターン認識の研究者がデータベース技術を理解し，データベースの研究者がパターン認識技術を理解することによって，革新的な技術開発のきっかけになるかも知れません．そのような協調の中で，何らかの貢献ができればと考えています．

■ 大規模ウェブアーカイブからのデータマイニング（豊田:DE）

□ スタンスと現状

Web は膨大な文書の集合であると同時にハイパーリンクで結合された文書の巨大なネットワークでもあります．Web ページは，生成，更新，および消滅の過程を経て日々変化しており，それに応じて Web のリンク構造も動的に変化を続けています．こうした Web の発展過程は実世界の動向と密接な関係を持つ傾向を強めつつあります．たとえば，テロのような重大な事件が発生すると，それに関した Web ページが次々と作成され，重要な情報には多くのページからハイパーリンクが張られていき，多数のユーザが訪れるようになっていきます．Web に現れる話題の大域的な構造を把握し，その変化を追跡することは，実社会に起きる事象の背景や予兆を探る上で有用と考えられますが，直近のスナップショットしか検索できない現状の検索エンジンでこれを達成することはできません．我々は，Web の過去から現在に至る履歴を蓄積する 10 数 TB 規模のアーカイブを構築し，様々な手法を用いて Web の空間構造および時系列変化の分析を行っています．

- Web の構造を俯瞰する地図の自動生成，および Yahoo!カテゴリとの比較 [11]
- Web の発展過程の抽出と可視化 [12]
- リンク解析を用いた Web スパムの調査

ウェブは，ページをノード，リンクをエッジとした巨大なグラフとして捉えられます．我々の最近のスナップショットでは，約 10 億ノード，40 億エッジ程度の規模となっています．このグラフは全体的には疎ですが，密な部分と疎な部分を併せ持つ複雑な構造をなしています．この規模のグラフになると，ノード数，エッジ数に対して $O(N^2)$ 以上の計算量を持つアルゴリズムを適用することは難しく，現状では，不要なノードおよびエッジの削減，線形に近いアルゴリズムによる近似など様々な工夫を要します．

□ パターン認識／抽出技術に期待するもの

極大クリークの列挙など，有用なグラフのパターンを抽出するアルゴリズムの計算量は大きく，巨大なウェブのグラフへ適用はできません．巨大なグラフにも適用可能なグラフ解析技術の開発に期待しています．また，動的なグラフにおける変化のパターン抽出については，まだ未開発な部分が多く，今後の発展が待たれます．

■ 編集映像からのパターン抽出とその利用（新田:PRMU）

□ スタンスと現状

映像編集とは，素材映像から選択した適切な部分映像を，音

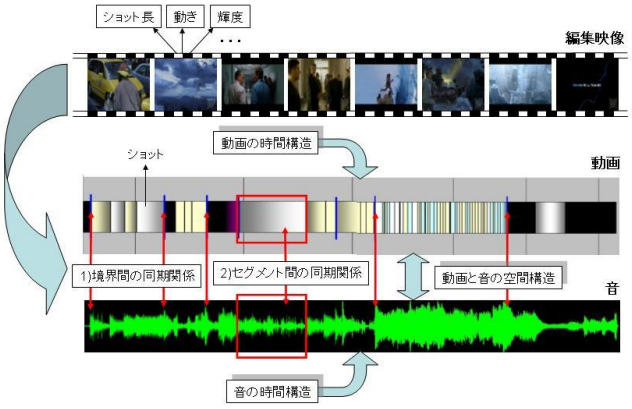

図 4　編集映像の感覚的空間・時間構造

響（音楽，効果音，ナレーションの音声など）やトランジション効果（カット，ディゾルブ，フェードイン／アウトなど）などを付与しながらつなぎ合わせ，新たな映像を制作する作業です．ここで，編集映像中の部分映像，トランジション効果，音響にはさまざまな種類が存在し，音響と画像などの複数メディア間の組み合わせや，各メディア内の時系列上での組み合わせが無限に存在するため，編集映像の質は編集者の持つ技術に依存します．そこで我々は，実際に専門家により製作された編集映像の解析により得られた編集パターンに基づいた映像編集支援を目的としています [14].

ここで，映像編集の目的として，ニュース番組中で放映されるスポーツのダイジェスト映像に代表される短時間での情報獲得を可能にする映像制作と，映画予告映像に代表される短時間で素材映像の魅力を伝えるエンタテイメント性の高い映像制作の 2 つが挙げられます．これまで前者に対し，映像解析による意味内容に基づいた重要場面の自動抽出などに取り組んできました [13]．しかし後者に対しては，映像の意味内容以外に，映像中の動きやカメラの切り替わりの頻度などの変化が生み出すリズムや，音の立ち上がりとカメラの切り替わりの同期関係が生み出す調和感など感覚的な側面について考慮する必要があります．そこで現在，後者に相当する映画予告映像を題材に，感覚的な側面に着目した以下のパターン抽出を試みています．

（1）音セグメントやショットといった各メディアを構成するセグメント間や，その境界間の関係といった，調和感を生み出す複数メディア間の同期パターン
（2）音セグメントやショットの，長さやセグメント内の変化に基づいた時間軸上の並び方といった，リズムを生み出す時系列パターン

映像は，音響や動画などの映像を構成する各メディアのフレーム，またはセグメントの並びを表す時系列データのセットと考えられます．両メディアのさまざまな特徴が互いに干渉しあい，リズム感や調和感を生成していると考えられますが，実際には，音響と動画のセグメント構造が異なったり，両メディア間で関連する特徴対が各フレームもしくはセグメントの局所的・大域的な前後関係に依存するなど，空間・時間構造を同時に考慮する必要があります．

□ パターン認識／抽出技術に期待するもの

実際には，映像の感覚的な側面には芸術性が大きく関わるため，映像の種類や，制作者など映像間の共通項により抽出するべきパターンは異なると考えられます．ユーザの要求に応じて，さまざまな映像の組み合わせでパターンを抽出するため，複数の時系列データセットから，空間・時間構造を同時に考慮しながら，同期パターン，時系列パターンを高速に抽出するような方法がないかと考えています．

文　　献

[1] 佐藤真一，中村裕一，“映像処理はもう古い？ そんなことはありません !? ～映像処理の学術的意義と将来展望～，” 信学技報 PRMU-2003-56，pp.77-80，2003.
[2] T. Asai, *et al.*, “Efficient substructure discovery from large semi-structured data,” Proc. 2nd SDM’02, SIAM, pp.158-174, 2002.
[3] S. Morinaga, *et al.*, “Key semantics extraction by dependency tree mining,” Proc. ACM KDD-05, pp.666-671, 2005.
[4] T. Uno, *et al.*, “LCM v.2: Efficient mining algorithms for frequent/closed/maximal itemsets,” Proc. 2nd IEEE ICDM Workshop FIMI’04, 2004.
[5] N. Katayama, H. Mo, I. Ide, and S. Satoh, “Mining Large-Scale Broadcast Video Archives towards Inter-Video Structuring,” PCM2004, LNCS 3332, pp.489-496, 2004.
[6] I. Ide, H. Mo, N. Katayama, and S. Satoh, “Topic threading for structuring a large-scale news video archive,” Proc. Third Intl. Conf. on Image and Video Retrieval, Lecture Notes in Computer Science, Vol.3115, pp.123-131, 2004.
[7] 尾関基行，伊藤雅嗣，里 雄二，中村裕一，大田友一，“複合コミュニティ空間における注目の共有 ～注目誘導行動による物体への注釈付け～，” 日本バーチャルリアリティ学会論文誌，Vol.8，No.4，pp.369-377，2003.
[8] 小泉敬寛，亀田能成，中村裕一，“個人行動記録の構造化類似検索手法，” 信学技報，Vol.PRMU2005-34，pp.1-6，2005.
[9] Y. Sakurai, M. Yoshikawa, S. Uemura, and H. Kojima, “Spatial indexing of high-dimensional data based on relative approximation,” VLDB Journal, Vol.11, No.2, pp.93-108, 2002.
[10] Y. Sakurai, S. Papadimitriou, C. Faloutsos, “BRAID: Stream Mining through Group Lag Correlations,” ACM SIGMOD Conference, pp.599-610, 2005.
[11] M. Toyoda and M. Kitsuregawa, “Creating a Web Community Chart for Navigating Related Communities,” Proc. of the Twelfth ACM Conf. on Hypertext and Hypermedia, pp.103-112, 2001.
[12] M. Toyoda and M. Kitsuregawa, “Extracting evolution of web communities from a series of web archives,” Proc. of the Fourteenth Conf. on Hypertext and Hypermedia, pp.28-37, 2003.
[13] N. Nitta, N. Babaguchi, and T. Kitahashi, “Generating Semantic Descriptions of Broadcasted Sports Videos Based on Structures of Sports Games and TV Programs,” Multimedia Tools and Applications, Vol.25, No.1, pp.59-83, 2005.
[14] 秋月達人，新田直子，馬場口登，“編集映像における動画と音ストリームのリズム同期について，” 第 4 回情報科学技術フォーラム，pp.233-234，2005.