

# 動画共有サイトにおける動画リストを用いた動画検索

西友規<sup>†</sup> 山口実靖<sup>†</sup>

動画共有サービスが普及し、それらの Web サイトには多くの動画がユーザによってアップロードされ、共有されている。多くの動画共有サイトですでに動画検索機能が提供されているが、動画共有サイトの普及と重要性の高まりを受け検索精度のさらなる向上は重要な課題の一つと考えることができる。本稿では、Web コミュニティの抽出手法と、ユーザが公開している動画リストの2点に着目し、これらにもとづく動画検索手法を提案する。本手法では、動画リストによる動画の登録を、Web コミュニティ抽出における Web ページから Web ページへのリンクと類似と考え、Web コミュニティ抽出手法を動画共有サイトに対して適用する。これにより動画共有サイトから動画コミュニティの抽出を行い、抽出された動画コミュニティを検索結果ととらえ動画検索を行う。提案手法を既存の検索手法と比較した結果、提案手法が既存手法より高い精度で動画の検索を行うことが可能であり、提案手法が有効であることが確認された。

## Video Search in Video Sharing Site Based on Web Community Extraction

Yuki Nishi<sup>†</sup> Saneyasu Yamaguchi<sup>†</sup>

A large number of video files are shared in video sharing services in the internet. In this paper, we propose a video community extraction method based on web community extraction methods, which consist of fan pages and center pages, and TF-IDF. With our proposed method, a video and a video list are considered as a center page and a fan page in web community extraction, respectively. For evaluation, we executed video search based on the proposed video community extraction and compared its search result with that of other methods. Our experimental results have shown that the proposed method have been able to provide better search results than other methods.

### 1. はじめに

インターネット上の動画共有サービスが普及し、多くの動画が動画共有サイトで共有されている。よって、動画共有サービス内における動画検索の重要性も向上していると考えられる。また、Web 空間からのコミュニティ抽出に関しては非常に多くの研究成果が得られており [1,2,3,4,5,6]、これらを動画共有サイト内の動画検索に応用することで、より良い検索を実現できると期待できる。

本稿では、まず動画共有サイトの機能である「タグ」と「動画リスト」について説明する。次に、既存研究である Web コミュニティの抽出手法と TF-IDF を紹介する。そして、Web コミュニティ抽出手法と TF-IDF を用いて動画コミュニティを抽出する手法と、それを用いた動画検索手法を提案する。最後に、評価実験の結果を示し提案手法の有効性を示す。

### 2. 動画共有サイトの機能

#### 2.1 タグ

多くの動画共有サイトでは、各動画の特徴を表す文字列をタグとして動画に対して付与することができる。例えばチャーハンの調理の動画なら、「チャーハン」や「料理」などのタグが付与されることと予想される。多くの場合、タ

グは動画の特徴を表しており、動画の検索、分類、説明などに利用される。タグを用いることにより、指定のタグが登録されている動画のみを検索したり、注目している動画と関連性のある動画を検索したりすることが可能となる。

#### 2.2 動画リスト

多くの動画共有サイトで、ユーザが指定した動画群を“動画リスト”として公開する機能が提供されている。動画リストは各ユーザが自由に作成することができるが、リスト内の動画同士には関連があることが多いと期待することができる。

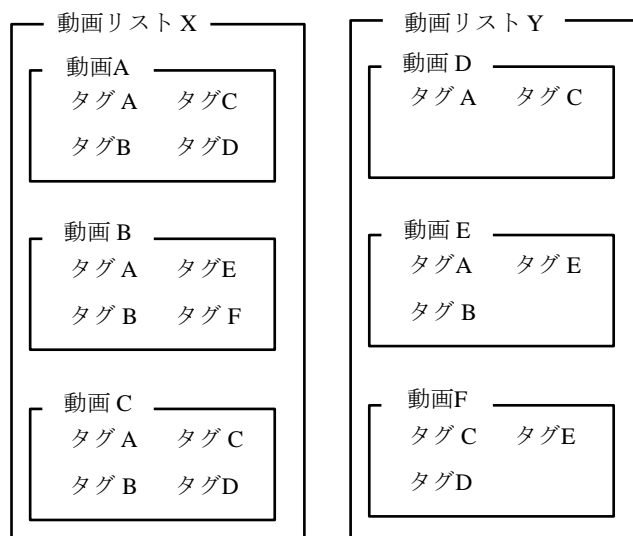


図 1 動画リスト

<sup>†</sup> 工学院大学 工学部 情報通信工学科  
Department of information and Communications Engineering, Kogakuin University

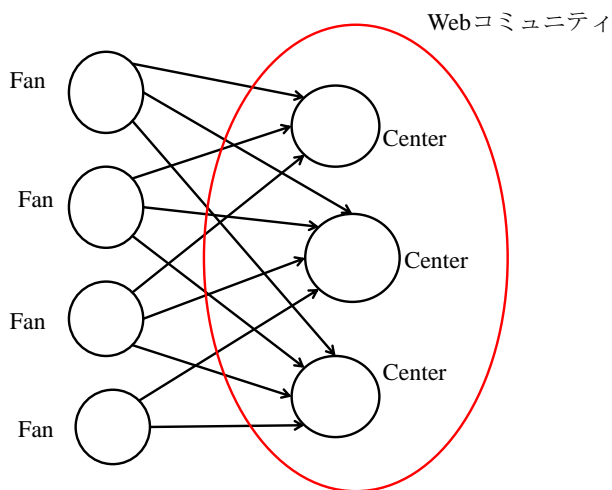


図 2 Center と Fan の関係

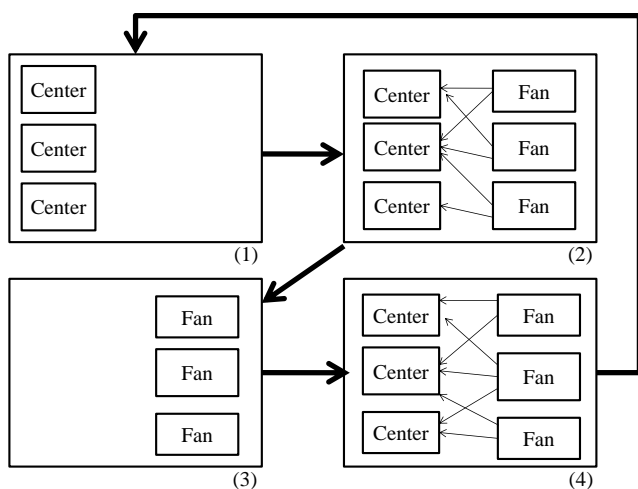


図 3 Web コミュニティ抽出手法

図 1 に動画リストと、各動画に付与されているタグのモデルを示す。各動画リストには 1 個以上の動画が登録されており、各動画にはその動画の特徴を表すタグが付与されている。

### 3. 既存研究

#### 3.1 Web コミュニティ抽出

Web 空間の中には共通の話題を有する Web ページ群が存在し、共通の話題を有するページ群を“Web コミュニティ”と呼ぶことができる。以下に、Web 空間からの Web コミュニティ抽出に関する研究を示す。

Kumar らは、Web ページ間のリンクグラフにおいて共通参照により構成される 2 部グラフが存在し、それらは共通の話題を持つとした。そして、密な 2 部グラフを Web コミュニティと定義した[1]。また、Web ページ群から完全 2 部グラフを抽出する手法を提案している。

Reddy らは、Web コミュニティを密な 2 部グラフ(DBG: Dense Bipartite Graph)とし、完全 2 部グラフでなく指定数以上のリンクを有する緩い共引用(Relax\_cocite)による 2 部グ

ラフの発見による Web コミュニティの抽出手法を提案している[2]。

村田は、Web の検索エンジンを用いてリンクの共起の完全 2 部グラフを抽出する手法を提案している[3]。

斉田らは、Web コミュニティ内のメンバが持つ話題が、シードの Web ページ群のトピックスから離れている量を表す距離量概念を加えて精度と網羅性を向上させる PlugDBS[4]を提案している。

これらの手法では、図 2 の様に特定の話題を持つ Web ページを Center とし、Center の集合を Web コミュニティと考える。また、Center の Web ページ群に対して多数のリンクを出している Web ページを Fan としている。この Center の集合と Fan の集合を用いて二部グラフを作成し、Center 集合からの Fan 集合の作成および Fan 集合からの Center 集合の作成を繰り返し Center 集合の改善を行っていく。

図 3 にこれらの手法における Web コミュニティ抽出の概要を示す。最初に初期 Center 集合あるいは初期の Fan 集合を与える。Center 集合を与えた場合、その Center 集合の内のすべてのページに対して、あるいは指定数以上のページに対してリンクを出しているページを Fan として抽出する。次に、この Fan 集合内のすべてのページから、あるいは指定数以上のページからリンクを受けるページを新しい Center として抽出する。以下同様に、Center 集合からの新しい Fan 集合の抽出、Fan 集合からの新しい Center 集合の抽出を繰り返すことで、ページ同士の関連が強い Center 集合を作成し、これを Web コミュニティとする。初期の Fan の与え方として、Kumar らは HITS[5]の HUB を応用する方法を示している。

また、Flake らは、集合内のページ間のリンク数が集合以外のページとのリンク数より多いページの集合を Web コミュニティとし、最大流問題を解くことにより Web コミュニティを抽出できることを示している。

#### 3.2 TF-IDF

TF-IDF は文書内の単語につける重みの一種であり、以下の TF と IDF により決定される。

TF (Term Frequency) は、文書内におけるその単語の出現頻度を表す。TF が大きい単語は文書内に頻繁に出現し、そのような単語は文書内において重要、代表的な単語であると考えられることができる。

IDF (Inverse Document Frequency) は、その単語を含む文書の頻度の逆数を表す。DF は文書頻度であり、その単語を含む文書の個数を表す。よって、DF が大きい単語は多くの文書に登場する一般的で重要度の低い単語と考えることができ、IDF が大きい単語が代表的な単語に適していると考えられることができる。

tfidf 値は、以下の式(1)の様に TF と IDF により計算される。tfidf 値が高い単語ほど、その文書にとって重要な単語と考えることができる。

$$\left. \begin{aligned} \text{tfidf} &= \text{tf} \times \text{idf} \\ \text{tf} &= \frac{n}{\max(n)} \\ \text{idf} &= \log\left(\frac{D}{d}\right) \end{aligned} \right\} \quad (1)$$

ただし、 $n$  は文書内におけるその単語の出現回数、 $\max(n)$  は文書内の全単語数、 $D$  は全文書数、 $d$  はその単語を含む文書数を表す。

#### 4. 提案手法

動画共有サイトにおいて「共通の話題を持つ動画の集合」を動画コミュニティとする。本章では、Web コミュニティ抽出手法を用いて動画コミュニティを抽出する手法(WC手法)を提案し、それを発展させた Web コミュニティ抽出手法と TF-IDF を併用して動画コミュニティを抽出する手法(WCTI 手法)を提案する。

##### 4.1 Web コミュニティ抽出手法を用いる動画コミュニティ抽出法 (WC 手法)

まず、Web コミュニティ抽出手法のみを用いる手法について述べる。本手法を WC 手法と呼ぶ。WC 手法では、Web コミュニティ抽出におけるリンク元ページ、リンク先ページ、リンクを、動画共有サイトにおける以下のものに置き換え、Web コミュニティ抽出手法を動画共有サイトに適用する。

- Web コミュニティ抽出における「リンク元ページ」を、動画共有サイトにおける「動画リスト」に置き換える。
- Web コミュニティ抽出における「リンク先ページ」を、動画共有サイトにおける「動画」に置き換える。
- Web コミュニティ抽出における「ページからページへのリンク」を、動画共有サイトにおける「動画リストによる動画の登録」に置き換える。

動画コミュニティ抽出において、Center は全動画集合から抽出し、Fan は全動画リストから抽出する。具体的には、Fan である動画リスト群から多くの登録を受けている動画を Center 動画とし、Center 動画群に対して多くの登録を行っている動画リストを Fan 動画リストとする。

そして図 4 の手順に従い動画コミュニティの抽出を行う。まず、人手により初期 Center 動画の集合を用意する。そして、Center 動画集合内の動画を指定個数以上登録している動画リストを Fan 動画リストとして抽出する。続いて、Fan 動画リスト集合内の指定個数以上の動画リストに登録されている動画を Center 動画として抽出する。以下同様に、Center 動画集合からの Fan 動画リストの抽出、Fan 動画リスト集合からの Center 動画の抽出を収束まで繰り返し、収束後の Center 動画集合を動画コミュニティとする。

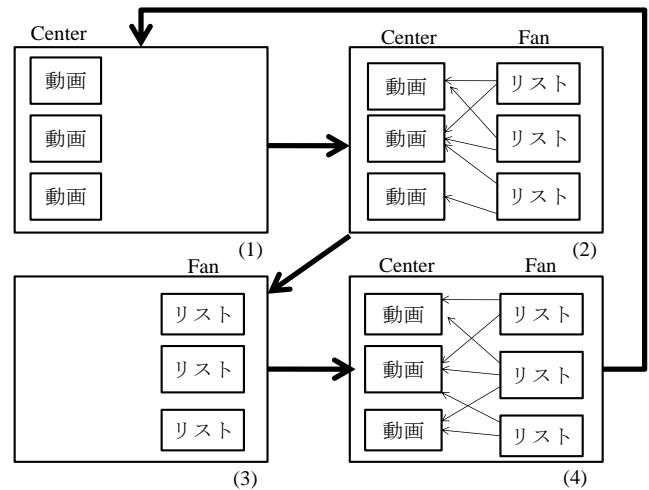


図 4 動画コミュニティ抽出法 (WC 手法)

##### 4.2 Web コミュニティ抽出手法と TF-IDF を用いる動画コミュニティ抽出法 (WCTI 手法)

次に、Web コミュニティ抽出手法と TF-IDF を用いる動画コミュニティ抽出法について述べる。本手法を WCTI 手法と呼ぶ。

本手法では、検索語が存在していることを前提に、検索語に関連する動画の動画コミュニティを抽出する。

前節の WC 手法同様に本手法でも、Web コミュニティ抽出におけるリンク元ページ、リンク先ページ、リンクを、動画共有サイトにおける動画リスト、動画、動画リストへの動画の登録で置き換え、Web コミュニティ抽出手法を適用する。

また、次項に示す方法で動画リスト内のタグに対して tfidf 値を定義し、Center 動画集合からの Fan 動画リストの抽出と Fan 動画リストの集合からの Center 動画の抽出の際にこの tfidf 値を用いて重み付けを行う。

##### 4.2.1 動画リストとタグにおける TF-IDF

図 1 の様に、動画リスト内には 1 個以上の動画が登録されており、それぞれの動画にはタグが登録されている。本稿では、動画リスト 1 個を文書 1 個、動画タグを単語、動画リストに含まれる全動画の全タグを文書内の全単語と見なし、3.2 節の式(1)を用いて動画リスト内のタグの tfidf 値を定義する。すなわち、 $n$  はその動画リスト内にそのタグが登場する回数、 $\max(n)$ はその動画リスト内の全タグ数、 $D$  は全動画リスト数、 $d$  はその単語を含む動画リストの数となる。

以下に、図 1 の例における、TF と IDF を示す。動画リスト X に存在するタグの数が 12 であるため、動画リスト X の全タグ数( $\max(n)$ 相当)は 12 となる。タグ C は動画リスト X に 2 回登場するため、タグ C の動画リスト X における TF 値は  $2/12$  となる。タグ A とタグ F の動画リスト X における TF 値は  $3/12$ ,  $1/12$  となる。また、動画リストは 2 個であるため、全動画リスト数(D 相当)は 2 なる。そして、タグ F は動画リスト X のみに登場するためタグ F の DF 値

は  $\log(1/2)$ , IDF は  $\log(2/1)$  となる。

#### 4.2.2 TF-IDF を考慮した Center 集合から Fan の抽出

WCTI 手法では、前項の手法に従い各動画リスト内における検索語の tfidf 値を計算し、これが高い動画リストを検索語に適した動画リストとみなす。具体的には、動画リスト  $l$  の評価を以下の式(2)の  $fti(l)$ により行い、 $fti(l)$ が高い 50 件の動画リストを Fan とする。

$$fti(l) = tfidf^{10} \times f(l) \quad (2)$$

ただし、 $f(l)$ は動画リスト  $l$  が含む Center 動画の数、 $tfidf$  は動画リスト  $l$  における検索語の tfidf 値、 $tfidf^{10}$  はその 10 乗である。10 乗とした理由は、 $f(l)$  に対して  $tfidf$  の影響を相対的に大きくするためである。

前節の WC 手法と本節の WCTI 手法の違いは、前節の WC 手法は  $f(l)$ のみで評価しており、本節の WCTI 手法は  $f(l)$  と  $tfidf$  の 10 乗の積で評価している点である。

#### 4.2.3 TF-IDF を考慮した Fan 集合から Center の抽出

Fan 動画リスト集合からの Center 動画の抽出においては、以下の式(3) の  $cti(v)$ により動画  $v$  を評価し、 $cti(v)$ が高い 50 件の動画を Center とする。

$$cti(v) = HasTag(v, t) + \sum_{l \in L} fti(l) \quad (3)$$

ただし、 $L$  は、「Center 動画を含んでいる動画リスト」の集合、 $HasTag(v, t)$ の値は動画  $v$  が検索語である  $t$  をタグに持てば 1, 持たなければ 0 である。

$\sum_{l \in L} fti(l)$ は、動画  $v$  が Fan 動画リストから含まれるごとに評価値  $fti(l)$ を与えられ、その評価値  $fti(l)$ は Center への関連度と検索語への関連度を加味したものとなっている。 $\sum_{l \in L} fti(l)$ は 1 と比べて十分に小さいため、実質的にはタグの有無( $HasTag$  が 1 であるか 0 であるか)により動画がクラス分けされ、同一クラス内における優先順位付けに  $\sum_{l \in L} fti(l)$ が使用されることとなる。

#### 4.2.4 WCTI 手法による動画コミュニティ抽出手順

WCTI 手法における動画コミュニティ抽出手順を図 5 および以下の(1)~(4)に示す。

(1) 共通の主題を持った動画を 10 件選択し、それを初期の Center 集合とする。

(2) 第 4.2.2 項の手法に従い、動画リストを  $fti(l)$ を用いて評価する。そして、値が高い動画リスト 100 件を Fan 動画リスト集合とする。

(3) 第 4.2.3 項の手法に従い、動画を  $cti(v)$ を用いて評価する。そして、値が高い動画 50 件を Center 動画集合とする。

(4) 収束をするまで、あるいは十分な回数、上記の(2)と(3)を繰り返す。

以上により得られた Center 集合を動画コミュニティとする。

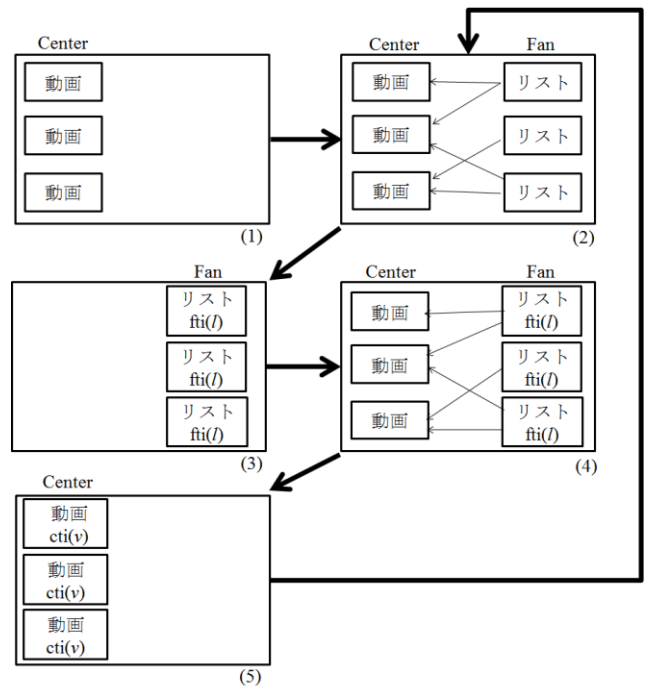


図 5 TF-IDF 法を用いる動画コミュニティ抽出法 (WCTI 手法)

## 5. 評価と考察

本章では、動画共有サイトで提供されている検索機能、提案手法(WC 手法)、提案手法(WCTI 手法)のそれぞれによる検索結果の比較を行う。

動画共有サイトにより提供されている検索手法の検索結果としては、フリーテキスト検索結果を再生回数順あるいは動画リスト登録回数順に並び替えて上位 50 件を検索結果としたもの、検索語をタグに含む動画群を再生回数順あるいは動画リスト登録回数順に並び替え上位 50 件を検索結果としたもの、の 4 通りを用いた。また比較のために、Web 検索エンジンにて検索範囲を該当動画共有サイトのみ指定し単語検索を行った結果も合わせて評価する。両提案手法では、抽出された動画コミュニティ内の動画の上位 50 件を検索結果とした。また初期 Center 動画の集合としては、動画共有サイトにより提供されているタグ検索の結果を動画リスト登録回数順に並び替えた上位 20 件から 10 件を手で選択したものを用いた。検索はクローラや評価が許されている実動画共有サービスを用いて行った。

検索結果の評価は人手で行い、評価者らが検索結果に含まれる各動画を再生、閲覧し主観により次の 3 段階の評価 (A, B, C)に分類した。

(A 評価) : 検索語と深い関係がある動画

(B 評価) : 検索語と関連があるが、関係が深くない動画

(C 評価) : 検索語と無関係の動画

検索語を政治家の名前(以下「政治家」としたとき、「チャーハン」としたとき、「世界遺産」としたとき、「MTG」としたときの第三者による評価結果を表 1 から表 8 に示

す。それぞれの表は一人の評価者により評価結果を表しており、評価者は著者とは異なる人間である。また、評価者に対しては全手法の検索結果に含まれる全動画の一覧のみが与えられ、どの動画がどの検索手法による検索結果であるかを評価者が特定できない状況で評価を行った。参考のために著者による評価結果も付録 A に示す。同様に、どの動画がどの検索手法による検索結果であるかを特定できない状況で評価を行った。

表 1 検索語「政治家」：第三者の評価結果(1)

	A	B	C
キーワード検索+再生数が多い順	17	16	17
キーワード検索+動画リスト登録数が多い順	13	16	21
タグ検索+再生数が多い順	17	16	17
タグ検索+動画リスト登録数が多い順	12	16	22
検索サイト(動画共有サイトのみを対象とする)	31	12	7
WC手法	12	16	22
WCTI手法	47	3	0

表 2 検索語「政治家」：第三者の評価結果(2)

	A	B	C
キーワード検索+再生数が多い順	12	29	9
キーワード検索+動画リスト登録数が多い順	16	28	6
タグ検索+再生数が多い順	12	30	8
タグ検索+動画リスト登録数が多い順	17	29	4
検索サイト(動画共有サイトのみを対象とする)	33	14	3
WC手法	12	25	13
WCTI手法	48	2	0

表 3 検索語「チャーハン」：第三者の評価結果(1)

	A	B	C
キーワード検索+再生数が多い順	5	0	45
キーワード検索+動画リスト登録数が多い順	1	0	49
タグ検索+再生数が多い順	26	5	19
タグ検索+動画リスト登録数が多い順	24	3	23
検索サイト(動画共有サイトのみを対象とする)	25	1	24
WC手法	0	0	50
WCTI手法	37	0	13

表 4 検索語「チャーハン」：第三者の評価結果(2)

	A	B	C
キーワード検索+再生数が多い順	4	3	43
キーワード検索+動画リスト登録数が多い順	2	1	47
タグ検索+再生数が多い順	25	9	16
タグ検索+動画リスト登録数が多い順	22	8	20
検索サイト(動画共有サイトのみを対象とする)	22	1	27
WC手法	1	2	47
WCTI手法	34	4	12

表 5 検索語「世界遺産」：第三者の評価結果(1)

	A	B	C
キーワード検索+再生数が多い順	2	3	45
キーワード検索+動画リスト登録数が多い順	4	4	42
タグ検索+再生数が多い順	15	8	27
タグ検索+動画リスト登録数が多い順	14	9	27
検索サイト(動画共有サイトのみを対象とする)	34	5	11
WC手法			
WCTI手法	40	5	5

表 6 検索語「世界遺産」：第三者の評価結果(2)

	A	B	C
キーワード検索+再生数が多い順	4	14	32
キーワード検索+動画リスト登録数が多い順	7	13	30
タグ検索+再生数が多い順	14	21	15
タグ検索+動画リスト登録数が多い順	19	16	15
検索サイト(動画共有サイトのみを対象とする)	27	11	12
WC手法			
WCTI手法	46	4	0

表 7 検索語「MTG」：第三者の評価結果(1)

	A	B	C
キーワード検索+再生数が多い順	8	15	27
キーワード検索+動画リスト登録数が多い順	4	13	33
タグ検索+再生数が多い順	8	15	27
タグ検索+動画リスト登録数が多い順	4	14	32
検索サイト(動画共有サイトのみを対象とする)	41	7	2
WC手法			
WCTI手法	46	1	3

表 8 検索語「MTG」：第三者の評価結果(2)

	A	B	C
キーワード検索+再生数が多い順	16	30	4
キーワード検索+動画リスト登録数が多い順	8	31	11
タグ検索+再生数が多い順	16	32	2
タグ検索+動画リスト登録数が多い順	9	36	5
検索サイト(動画共有サイトのみを対象とする)	42	7	1
WC手法			
WCTI手法	46	4	0

すべての評価結果において、Web コミュニティ抽出手法と TF-IDF を併用する WCTI 手法が最も(A 評価)が多く、最も(C 評価)が少なくなり、本提案手法が有効であることが分かった。また、単純に Web コミュニティ抽出手法を用いたのみの WC 手法は必ずしも優れた検索結果を提供できないことも分かった。

WC 手法では、初期の Center 集合として検索語と適合度が高い動画群を与えても、Fan の抽出と Center の抽出を繰り返すに従い検索語を含むより抽象的な概念の動画の集合と変わって行ってしまった。たとえば、初期の Center に「チャーハン」の動画群を与えた場合は、「チャーハン」より抽象度が高い「料理」に関する動画の集合に収束した。

WC 手法と WCTI 手法の比較より、動画リストには話題の選定に有効なものもあるが、有効でないものも含まれており、検索にはこれらの区別が重要であると考えられる。

検索語が「世界遺産」と「MTG」である場合、WC 手法により抽出された動画コミュニティのサイズが動画 0 件、5 件となり、50 件の検索結果を提供できず比較不可能となった。

また、WCTI 手法において  $f_{ti}(l)$  を、 $tfidf^{10}$  を用いる式(2)ではなく下記の  $tfidf$  の 1 乗を用いる式(4)により計算した場合の評価結果を付録 B に示す。

$$f_{ti}(l) = tfidf \times f(l) \quad (4)$$

## 6. おわりに

本稿では、動画コミュニティ抽出法と TF-IDF を使用した動画検索手法を提案した。提案手法と他の検索手法の検



索結果を評価した結果、提案手法は他の検索手法に比べて検索語と関連の高い動画をより多く抽出可能であることが確認され、有効性が示された。今後は、さらに多くの検索語による評価をし、さらに精度を向上させるための方法を考察する予定である。

## 謝辞

本研究はJSPS 科研費 22700039の助成を受けたものである。

検索結果の評価のために、非常に多くの動画の閲覧を行った評価者の人たちに感謝の意を表す。

## 参考文献

- 1) Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, Andrew Tomkins, "Trawling the Web for emerging cyber communities," In Proc. of the 8th international conference on World Wide Web, pp. 1481 - 1493, 1999.
- 2) P. Krishna Reddy, Masaru Kitsuregawa, "An approach to relate the web communities through bipartitegraphs," Proc. of the 2nd International Conference on Web Information Systems Engineering, 2001.
- 3) 村田 剛志, "参照の共起性に基づく Web コミュニティの発見", 人工知能学会論文誌, Vol. 16, No. 3, pp. 316-323, 2001.
- 4) Naoyuki Saida, Akira Umezawa, Hayato Yamana, "PlusDBG: Web Community Extraction Scheme Improving Both Precision and Pseudo-Recall", In Proc. of the 7th Asia-Pacific Web Conference, 2005.
- 5) Jon M. Kleinberg, "Authoritative sources in a hyperlinked environment," Journal of the ACM (JACM), Volume 46 Issue 5, pp. 604 - 632, 1999.
- 6) Gary William Flake, Steve Lawrence, C. Lee Giles, "Efficient identification of Web communities," In Proc. of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 150-160, 2000.

## 付録

### A. 著者による評価

著者による評価結果を以下に示す。同様に提案手法のWCTI手法が最も優れた結果となった。

表 9 検索語「政治家 A」：著者の評価結果

	A	B	C
キーワード検索+再生数が多い順	17	20	13
キーワード検索+動画リスト登録数が多い順	15	18	17
タグ検索+再生数が多い順	18	20	12
タグ検索+動画リスト登録数が多い順	14	20	16
検索サイト(動画共有サイトのみを対象とする)	33	12	5
WC手法	14	18	18
WCTI手法	48	2	0

表 10 検索語「チャーハン」：著者の評価結果

	A	B	C
キーワード検索+再生数が多い順	6	0	44
キーワード検索+動画リスト登録数が多い順	2	0	48
タグ検索+再生数が多い順	28	5	17
タグ検索+動画リスト登録数が多い順	23	4	23
検索サイト(動画共有サイトのみを対象とする)	24	0	26
WC手法	1	0	49
WCTI手法	37	2	11

表 11 検索語「世界遺産」：著者の評価結果

	A	B	C
キーワード検索+再生数が多い順	2	7	41
キーワード検索+動画リスト登録数が多い順	4	9	37
タグ検索+再生数が多い順	11	19	20
タグ検索+動画リスト登録数が多い順	13	14	23
検索サイト(動画共有サイトのみを対象とする)	24	14	12
WC手法			
WCTI手法	40	9	1

表 12 検索語「MTG」：著者の評価結果

	A	B	C
キーワード検索+再生数が多い順	11	36	3
キーワード検索+動画リスト登録数が多い順	6	36	8
タグ検索+再生数が多い順	11	38	1
タグ検索+動画リスト登録数が多い順	7	40	3
検索サイト(動画共有サイトのみを対象とする)	39	10	1
WC手法			
WCTI手法	46	2	2

### B. tfidf の 1 乗を用いる手法

式(2)ではなく式(4)を用いる WCTI 手法による検索結果の評価を以下に示す。式(4)を用いる WCTI 手法では十分な検索精度が得られないことが分かる

表 13 検索語「政治家」：著者の評価結果

	A	B	C
式(4)を用いるWCTI手法	11	34	5