

プラズマ動画の類似部分検索のための 3 次元周波数を用いた非類似度

3D frequency-based dissimilarity for partial similarity retrieval of plasma movies

武田 佳和[†] 宝珍 輝尚[†] 野宮 浩揮[†] 中西 秀哉^{††} 小嶋 護^{††}

Yoshikazu Takeda Teruhisa Hochin Hiroki Nomiya Hideya Nakanishi Mamoru Kojima

1. はじめに

核融合科学研究所等で行われている核融合実験では、熱量をセンサーで計測したデータやプラズマ外部の磁場の変動（磁場揺動）を計測したデータなどの、大量の、しかも、さまざまな測定データが得られる [1]。核融合科学研究所のプラズマ実験において撮影されるプラズマ発光動画もその一つである。

これらの動画データはストレージに蓄積されており、蓄積された大量の動画データから、最近発見された興味深い現象の動画データと類似のものを取り出したいという要求がある。しかし、現状では、実験パラメータの類似性に基づいた検索が行われる程度である。実験パラメータが異なっても類似のものがある可能性があり、また、このようなものこそが新たな発見に結びつく可能性があると考えられるのであるが、蓄積されている大量の動画データに対して目視により類似性を判定することは事実上不可能であり、計算機を利用した支援が望まれている。

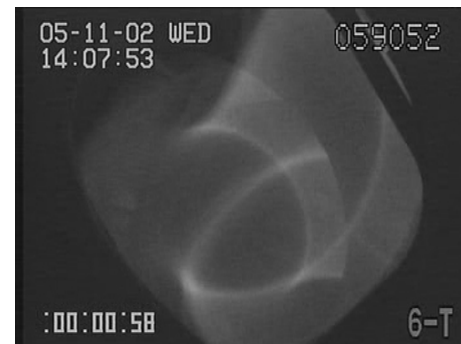
これまでに、動画データを一定フレーム数の区間に区切り、その区間に対して 3 次元フーリエ変換を行い、得られたフーリエ係数を使用して動画の部分検索を行う方法について検討してきた [2]。ここでは、情報量が多いと考えられる、分散や変動係数（= 分散 / 平均）の大きい 10 個のフーリエ係数を用いて非類似度を計算している。しかし、検索実験を行ったところ、あまり良い精度は得られなかった。これは、使用している非類似度が動画データの非類似性をよく反映していなかったからではないかと考えられる。

そこで本研究では、核融合科学研究所の核融合実験で撮影されている多くのプラズマショット動画から、有用となるデータを抽出することを目的として、動画の一部を検索キーとする類似部分検索に用いる非類似度を提案する。提案する非類似度は、3 次元フーリエ変換を使用し、人間の視覚特性を考慮して非類似度を計算するものである。

以下、2. で研究対象とする動画について述べ、3. で 3 次元フーリエ変換について述べる。4. で部分検索手法を概説し、5. で非類似度を提案する。6. で、提案した非類似度の良好性を評価する検索実験について述べ、最後に 7. でまとめる。

2. 研究対象とする動画

本研究では、核融合科学研究所で行われた実験データの中で、プラズマショットを撮影した動画を用いる [2]。プラズマショットの例を図 1 に示す。



(a) 例 1



(b) 例 2



(c) 例 3

図 1 プラズマショット動画

[†] 京都工芸繊維大学, Kyoto Institute of Technology

^{††} 核融合科学研究所, National Institute for Fusion Science

プラズマ発光動画は 29.97[frame/sec] の MPEG-1 で符号化されている．図 1 を見ると動画の中にはその撮影時の情報が付加されているのが分かる．動画の左上にはその撮影時の日付，右上にはショット番号，左下にはプラズマ生成からの経過時間，そして右下にはカメラ位置の識別子が記載されている．

このショット例は，2005 年 11 月 2 日に行われたショット番号 59052 番の動画の，撮影開始から 0.58 秒 (図 1(a))，1.18 秒 (図 1(b))，2.35 秒 (図 1(c)) のときの様子である．また撮影は 6-T という場所に設置されたカメラで行われており，これは可視光を撮影するカメラである．

これら画面中の文字は人が動画を区別するのには便利であるが，動画像のみを用いて行う類似部分検索にはノイズになると考えられる．よって本研究では，動画のアスペクト比 (縦横比) を保った上で画面の中心部分を切り抜いて用いる．具体的には，画面を 320[pixel]×240[pixel] の大きさにリサイズした後，高速フーリエ変換 (FFT) を用いるために中心の 256[pixel]×128[pixel] 画素の部分をトリムした．また，動画は輝度成分 (8bit) のみを扱っている．

3. 離散フーリエ変換

2. で説明した動画を対象に，周波数成分を取り出すために 3 次元フーリエ変換を用いる [3]．動画を輝度分布 $x(m, n, t)$ ， ($m, n = \text{pixel}$, $t = \text{frame}$) とみなし， x を離散フーリエ変換 (DFT) した周波数成分 $X(u, v, f)$ は式 (1) から式 (3) のように表される．

$$\begin{aligned} X(u, v, f) &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{t=0}^{T-1} x(m, n, t) \\ &\cdot \exp \left\{ -2\pi i \left(\frac{mu}{M} + \frac{nv}{N} + \frac{tf}{T} \right) \right\} \\ &= |X(u, v, f)| \exp \{ i\phi(u, v, f) \} \end{aligned} \quad (1)$$

$$|X(u, v, f)| = \sqrt{R^2(u, v, f) + I^2(u, v, f)} \quad (2)$$

$$\phi(u, v, f) = \tan^{-1} \frac{I(u, v, f)}{R(u, v, f)} \quad (3)$$

ただし u, v, f は周波数変数， i は虚数単位， M, N は動画の width および height， T は時間成分の長さ， $R(u, v, f), I(u, v, f)$ は $X(u, v, f)$ の実数部および虚数部， $|X(u, v, f)|$ は振幅成分， $\phi(u, v, f)$ は位相成分である．

また式 (4) の操作により逆変換可能で，もとの $x(m, n, t)$ に復元できる．

$$\begin{aligned} x(m, n, t) &= \sum_{f=0}^{T-1} \sum_{v=0}^{N-1} \sum_{u=0}^{M-1} X(u, v, f) \\ &\cdot \exp \left\{ 2\pi i \left(\frac{ft}{T} + \frac{vn}{N} + \frac{um}{M} \right) \right\} \end{aligned} \quad (4)$$

4. 部分検索

本研究では，部分検索の方法として，動画を一定フレーム数の区間に区切り，その区間ごとに非類似度を計算して比較していく方法をとる [2, 4]．

一区間中のフレーム数は 16 フレームとした．これは，30fps の動画では約 0.5 秒であり，区間内でプラズマの変化がほどよく収まる秒数となっている．従って，使用する 3 次元フーリエ変換の次元は $\text{width} \times \text{height} \times \text{depth} = 256 \times 128 \times 16$ となる．ただし，実際に用いるのは対称性から各次元の半分までである．

5. 非類似度

5.1 周波数と視認性

本論文で提案する非類似度は，周波数における人の目の特性を考慮して作成したものである．人の目は，周波数が高くなるにつれて，位相差および周波数の違いが分かりにくくなるということが分かっている [4, 5]．具体的には次のような視認特性がある．

- 低周波域では，同一周波数における位相差が判別可能．
- 中周波域では，隣接する周波数の違いが判別可能．
- 高周波域では，ほとんど違いが判らない．

5.2 3次元周波数を用いた非類似度

5.1 で述べた特性を考慮した非類似度 D を式 (5) に示す．ここで，検索対象の 3 次元周波数を $X = X(f_x, f_y, f_t)$ とし，検索キーの 3 次元周波数を $Y = Y(f_x, f_y, f_t)$ とする．

$$\begin{aligned} D &= C_{DC} |X_{DC} - Y_{DC}| \\ &+ C_{l0} \sqrt{\sum_{f_y=0}^{f_{y1}(0)} \sum_{f_x=\text{lowbeg}(0, f_y)}^{\text{low}(0, f_y)} |X - Y|^2} \\ &+ C_{lN} \sqrt{\sum_{f_t=1}^{f_{t1}} \sum_{f_y=0}^{f_{y1}(f_t)} \sum_{f_x=\text{lowbeg}(f_t, f_y)}^{\text{low}(f_t, f_y)} |X - Y|^2} \\ &+ C_{m0} \sqrt{\sum_{f_y=0}^{f_{ym}(0)} \sum_{f_x=\text{low}(0, f_y)+1}^{\text{mid}(0, f_y)} (|X| - |Y|)^2} \\ &+ C_{mN} \sqrt{\sum_{f_t=1}^{f_{tm}} \sum_{f_y=0}^{f_{ym}(f_t)} \sum_{f_x=\text{low}(f_t, f_y)+1}^{\text{mid}(f_t, f_y)} (|X| - |Y|)^2} \\ &+ C_{h0} \sqrt{\left(\sum_{f_y=0}^{f_{yh}(0)} \sum_{f_x=\text{low}(0, f_y)+1}^{f_{xh}} |X| - |Y| \right)^2} \\ &+ C_{hN} \sqrt{\left(\sum_{f_t=1}^{f_{th}} \sum_{f_y=0}^{f_{yh}(f_t)} \sum_{f_x=\text{low}(f_t, f_y)+1}^{f_{xh}} |X| - |Y| \right)^2} \end{aligned} \quad (5)$$

第 1 項は DC 成分 ($(f_x, f_y, f_t) = (0, 0, 0)$) の項であり， C_{DC} はその係数である．DC 成分は区間を通しての動画の明るさを意味しており，動画の類似検索においては用いたほうがよいと考えられる．

第 2 項および第 3 項は低周波域の項である．まず第 2 項は $f_t = 0$ での低周波域の周波数に対して，位相差を考慮した計算 (複素数計算) を行って足し込んだものである．第 3 項は $f_t \geq 1$ での低周波域の周波数に対して第 2 項と同様の計算を行ったものである．ここで， C_{l0} は $f_t = 0$ での低周波域への

係数, C_{lN} は $f_t \geq 1$ での係数である。また, $\text{lowbeg}(f_t, f_y)$ は f_t, f_y における低周波域の f_x の最小値を返す関数であり, $\text{low}(f_t, f_y)$ は f_t, f_y における低周波域の f_x の最大値を返す関数である。 $\text{fyl}(f_t)$ は f_t における低周波域の f_y の最大値を返す関数であり, ftl は低周波域での f_t の最大値を示す。

第4および第5項は中周波域の項である。これらは中周波域の周波数に対して、位相差の情報を用いない、絶対値の差の二乗を計算して足し込んだものである。 $\text{fym}(f_t)$ は $\text{fyl}(f_t)$ の中周波域版であり, $\text{mid}(f_t, f_y)$ は $\text{low}(f_t, f_y)$ の中周波域版である。 ftm と C_{m0} , さらに C_{mN} も同様に ftl , C_{l0} , C_{lN} の中周波域版である。

第6および第7項は高周波域の項である。これらは高周波域の周波数に対して、その絶対値の差をまとめて足し込んだものである。中周波域との違いは、絶対値の差について2乗を行っておらず、おおまかな足し算を行っていることである。ここでも C_{h0} , C_{hN} はそれぞれ高周波域における $f_t = 0$, $f_t \geq 1$ での係数であり, $\text{fyh}(f_t)$ は $\text{fyl}(f_t)$ の高周波版, fth は ftl の高周波版であり, fxh は高周波での f_x の終わり値である。

5.3 各帯域の決定

5.2 で示した非類似度を実際に計算するには, ftl や $\text{low}(f_t, f_y)$ 等が返す値を決定しなければならない。これはつまり, 5.1 で説明した視認特性を満たす範囲を決定することになる。

5.3.1 決定方法

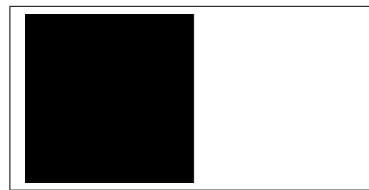
次に示すような一区間分の動画(約0.5秒)を作成し, 5.1 で述べた条件を満たすような周波数の範囲を, 実際に目で見ることによって決定する。

- 画面サイズは, 256×128 である。フレーム数は, 約一区間分である15フレームで, 29.97fps である。
- 各次元の周波数に対応するように, 画面および時間方向を分割する。

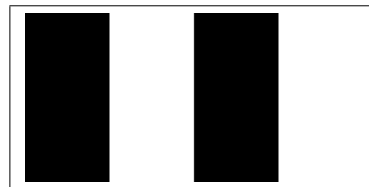
実際に画面を分割した画像を図2に示す。

図2(a)では, f_x 方向に1回, f_y 方向に0回(つまり $f_x = 1, f_y = 0$)の分割を行っており, 左上端は黒から始まっている。また, 図2(b)では, $f_x = 2, f_y = 0$ の分割を行っており, 左上端は黒から始まっている。図2(c)は, 図2(b)と似ているが, 左上端が白から始まっている。最後に図2(d)は, $f_x = 2, f_y = 2$ で分割し, 左上端が黒から始まるものである。

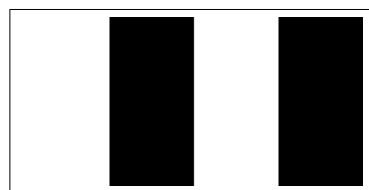
時間成分を分割しない場合($f_t = 0$)には, 図2のような画像を15フレーム連続させる(例: 分割が $f_t = 0, f_x = 1, f_y = 0$ の場合, 図2(a)を15フレーム連続させて動画とする)。この時, 「位相差がわかる」とは, 画面分割は同じである, 黒始まりの動画と白始まりの動画を順次見て, 違いがわかるということである(例: 図2(b)による動画と図2(c)による動画を見分けられる)。また, 「隣接する周波数の違いがわかる」とは, ある次元の周波数が1異なる動画(左上端の色は同じ)を順次見て, 違いがわかるということである(例: 図2(a)による動画と図2(b)による動画を見分けられる)。



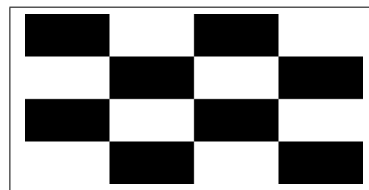
(a) $f_x = 1, f_y = 0$ に分割, 黒始まり



(b) $f_x = 2, f_y = 0$ に分割, 黒始まり



(c) $f_x = 2, f_y = 0$ に分割, 白始まり



(d) $f_x = 2, f_y = 2$ に分割, 黒始まり

図2 画面を分割した例(各画像は枠で囲っている)

時間成分を分割する場合, 例えば $f_t = 1, f_x = 2, f_y = 0$ 黒始まりの場合には, 図2(b)を8フレーム, 図2(c)を7フレーム連続させて動画とする。また $f_t = 1, f_x = 2, f_y = 0$, 白始まりの場合には, 図2(c)を8フレーム, 図2(b)を7フレーム連続させて動画とする。この時の位相差がわかるとは, これら二つを見て差がわかるということである。次に, $f_t = 2, f_x = 2, f_y = 0$, 黒始まりの場合には, 図2(b)を4フレーム, 図2(c)を4フレーム, 図2(b)を4フレーム, 図2(c)を3フレームと連続させて動画とする。この時の隣接する周波数の違いがわかるとは, $f_t = 1, f_x = 2, f_y = 0$ による動画と $f_t = 2, f_x = 2, f_y = 0$ (左上端は同じ)の二つを見て差がわかるということである。

以上のような動画を著者の一人が見て, 差がわかるかを判断した。

5.3.2 周波数帯域

5.3.1 で示した実験の結果, 次のように帯域を決定した。

低周波域 $f_t = 0$ の場合は図3のLowで示す点である。 $f_t = 1$ の場合は図4のLowで示す点である。これ以上の f_t では低周波域は存在しない。

中周波域 $f_t = 0$ の場合は図3のmiddleで示す点である。

$f_t = 1$ の場合は図 4 の Low 以外の範囲全てである ($f_x < 128, f_y < 64$) . $2 \leq f_t \leq 4$ では全範囲 ($0 \leq f_x < 128, 0 \leq f_y < 64$) である . $5 \leq f_t$ では存在しない .

高周波域 $f_t = 0$ の場合は図 3 の DC, Low, middle 以外の範囲全てである . $2 \leq f_t \leq 4$ では存在しない . $5 \leq f_t < 8$ では全範囲が高周波となる .

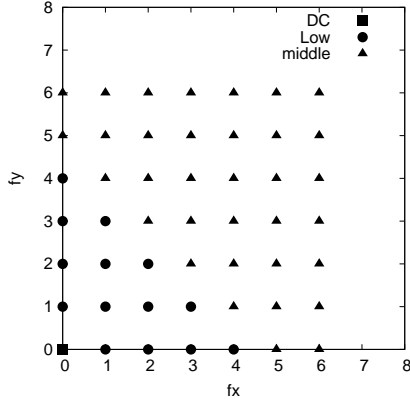


図 3 時空間周波数帯域 ($f_t = 0$)

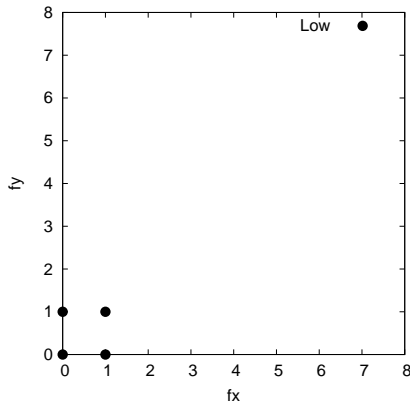


図 4 時空間周波数帯域 ($f_t = 1$)

この帯域を返すように $\text{low}(f_t, f_y)$ 等を設計した .

6. 検索実験

5.2 で説明した非類似度の性能を測るために検索実験を行う . 検索対象としては 10 個のプラズマショット動画を用いる . これらを区間に分割すると 120 区間となった . 以後, プラズマショット番号とそのセグメント(区間)を組み合わせて「ショット番号-セグメント」と表記する (例: ショット番号 40810, セグメント番号 0 の場合 40810-0) .

また, 検索キーは 40215-3 と 40800-6 の 2 個を用いる . これらのキーに対する正解ショット-セグメントは, 著者のうち 2 人に目視で検索対象を確認してもらい, 少なくとも 1 人が似ているとしたものである .

それぞれの検索キーについては, 適合率と再現率を用いて評価する . 適合率は式 (6) で, 再現率は式 (7) で求められる .

$$\text{適合率 (precision)} = \frac{\text{検索された正解集合の要素数}}{\text{検索した総数}} \quad (6)$$

$$\text{再現率 (recall)} = \frac{\text{検索された正解集合の要素数}}{\text{正解集合の全要素数}} \quad (7)$$

6.1 検索キー: 40215-3 での結果

非類似度における係数を $C_{DC} = C_{l0} = C_{lN} = C_{m0} = C_{mN} = C_{h0} = C_{hN} = 1$ と置いた場合を normal とする . また, 各帯域の計算結果より, 検索精度が向上するように設定した係数 $C_{DC} = 0.5, C_{l0} = 0.8, C_{lN} = 0.9, C_{m0} = 0.8, C_{mN} = 0.9, C_{h0} = 0, C_{hN} = 0.9$ を customize とする . これらの適合率・再現率曲線を図 5 に示す . 図 5 からわかるように, 係数を変えることにより若干性能が上がっている . その向上の最大の原因は, C_{h0} を 0 としたことである . つまり, C_{h0} に係わる項は類似性にはノイズであったのだと考えられる . これは, C_{h0} の範囲である $f_t = 0$ での高周波成分の数が多く, 目では区別できないようなものまで非類似度として算出してしまったために, 検索結果に悪影響を及ぼしていたのだと考えられる .

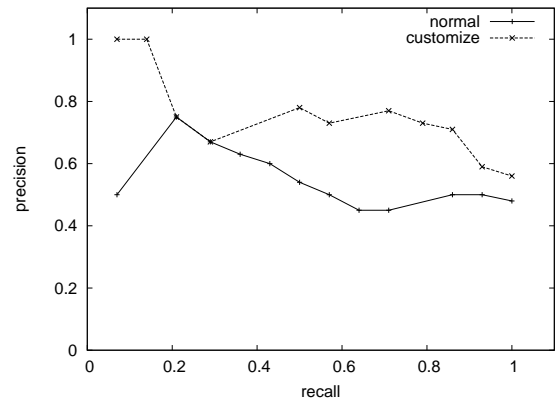
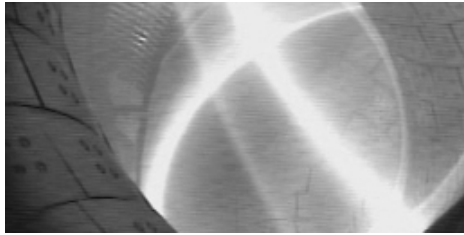


図 5 キー: 40215-3 での適合率・再現率曲線

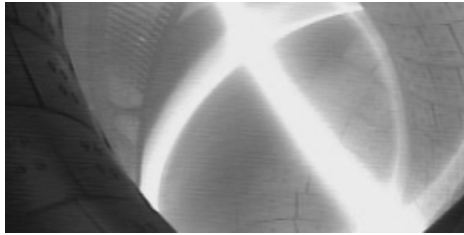
ここで, 検索キーのサムネイルと, 検索結果の上位に現れた区間のサムネイルを図 6 に示す . 図 6(b) の正解区間を目で確認すると, 光の推移と光の筋が類似していた . 次に, 図 6(c) の不正解区間を目で確認すると, 光り方の推移は似ているが, 図でわかるように光の中の筋が異なっていた . 従って, 光の筋の違いが区別できていないと考えられる .

6.2 検索キー: 40800-6 での結果

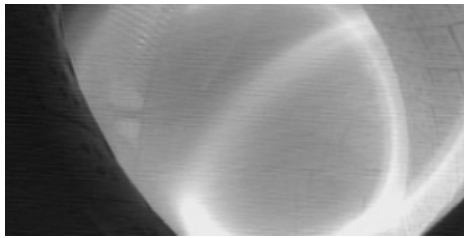
前節と同様に, normal と customize を用意した . customize は $C_{DC} = 0.01, C_{l0} = 0.5, C_{lN} = 1, C_{m0} = 0.3, C_{mN} = 0.7, C_{h0} = 0, C_{hN} = 0.06$ としたものである . これらの適合率・再現率曲線を図 7 に示す . 前節と同じく, 係数を変えることにより若干の性能改善がみられる . この最大の原因は, やはり C_{h0} を 0 としたことである . また, C_{lN} 以外の項を下げることによって性能が上がっている . これは, このキーでは低周波の動き成分 ($f_t \neq 0$) が重要であったのだと考えられる .



(a) キー：40215-3 のサムネイル



(b) 検索結果：正解区間 (40995-3) のサムネイル (customize で 1 位)



(c) 検索結果：不正解区間 (40830-3) のサムネイル (customize で 3 位)

図 6 キー：40215-3 のサムネイルと、検索結果の上位区間のサムネイル

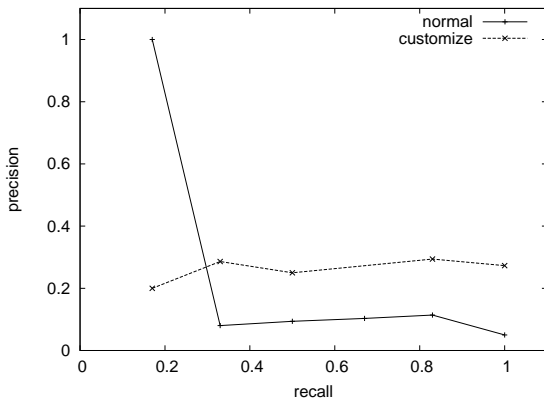
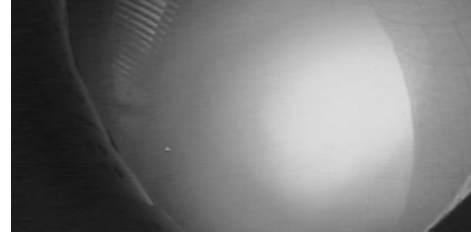


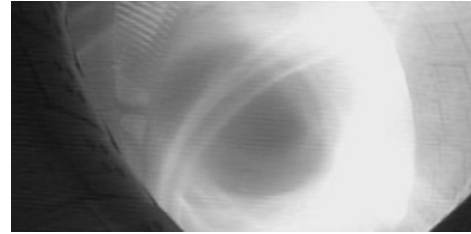
図 7 キー：40800-6 での適合率・再現率曲線



(a) キー：40800-6 のサムネイル



(b) 検索結果：正解区間 (40830-7) のサムネイル (customize で 5 位)



(c) 検索結果：不正解区間 (41068-8) のサムネイル (customize で 2 位)

図 8 キー：40800-6 のサムネイルと、検索結果の上位区間のサムネイル

次に、検索キーのサムネイルと、検索結果の上位に現れた区間のサムネイルを図 8 に示す。図 8(b) の正解区間では、画面全体にぼやけた光が覆っており、その光り方も類似していた。次に、図 8(c) の不正解区間を目で確認すると、図のように模様が入ってしまっており、その区別が付いていないと考えられる。

7. おわりに

本論文では、プラズマ動画の類似部分検索に関する非類似度を提案した。これは、人間の感覚に沿った非類似度を目指して作成したものである。実験の結果、係数を変えることにより性能が変動することを示し、検索キーにあった係数にすることで性能を向上できることを示した。

今後の課題としては、どのようなプラズマ動画に対して、どのように係数を決めればいいのかを判断できるようにすること、および、光の模様などの細部を区別できるようにすることがある。また、各帯域の分け方や非類似度の計算法についても、さらに検討する余地がある。

謝辞

本研究は、一部、NIFS 共同研究 NIFS09KCHH025 により支援されている。

参考文献

- [1] 宝珍輝尚, 中西秀哉, 小嶋護, 野宮浩揮, “核融合実験データと類似波形検索”, 電子情報通信学会技術研究報告 DE2009-11, vol.109, no.186, pp.1-6, 2009.
- [2] 武田佳和, 宝珍輝尚, 野宮浩揮, 中西秀哉, 小嶋護, “3次元 FFT を用いたプラズマ動画の類似部分検索”, DEIM フォーラム 2010 論文集, F1-3, 2010.
- [3] 山名信弘, 井辺昭人, 三浦文裕, 前島謙宣, 森島繁生, “動画の 3次元周波数成分を用いた顔認証システム”, 電子情報通信学会技術研究報告, vol.106, no.75, pp.13-18, 2006.
- [4] 山内祥祐, 宝珍輝尚, 野宮浩揮, 中西秀哉, 小嶋護, “揺動型時系列データに対する高速類似部分検索”, 情報処理学会研究報告 2009-DBS-148(16), 2009-F1-95, pp.1-8, 2009.
- [5] T. Hochin, K. Koyama, H. Nakanishi, M. Kojima, “Extension of frequency-based dissimilarity for retrieving similar plasma waveforms,” Fusion Engineering and Design, 83 (2), pp.417-420, 2008.