

複数デバイスで記録されたコンテンツの分類・同期システムの提案

石川 憲一[†] 森嶋 厚行[†] 杉本 重雄[†]

本論文ではデジタルカメラやデジタルビデオカメラといったデジタルデバイスで異なる形式で記録されたコンテンツを「イベント」に基づいて分類し同期を行うためのシステムを提案する。ここでいうイベントとは、例えば会議や旅行といったコンテンツの分類の手がかりとなる出来事の事である。本システムが扱うコンテンツの具体的な例としては、画像や映像、音声ファイルなどを対象とする。本論文では、このようなシステムの構想と実現のための要素技術について議論する。

Proposal of a System for Classification and Synchronization of Contents Recorded by Multiple Devices

KEN'ICHI ISHIKAWA,[†] ATSUYUKI MORISHIMA[†]
and SHIGEO SUGIMOTO [†]

This paper proposes a system to classify and synchronize contents recorded by digital devices such as digital cameras or digital video cameras based on the concept of "events." Here, events are something like meetings and travels, for classifying contents. The contents the system deals with include images, movies, and audios. This paper discusses its framework and technologies to realize such a system.

1. はじめに

近年、デジタルカメラやデジタルビデオカメラ、音声レコーダなどのデジタルデバイスが普及しつつある。また、記憶装置の大容量化と低価格化によって画像や映像、音声といったデジタルデバイスによって記録された大量のデジタルコンテンツが蓄積されるようになってきている。このように蓄積されたデジタルコンテンツの分類は多くの場合手動で行われることがほとんどである。コンテンツの量が少ない場合は人手による分類も可能であろうが、分類するコンテンツの量が膨大であったり、形式の異なるコンテンツ(例えば画像と動画)が混在している場合、中身を逐一確認しなければならず現実的とは言えない。

コンテンツの分類や検索の際重要となるのがメタデータである。例えばデジタルカメラで撮影した画像には Exif(Exchangeable Image File Format)¹⁾ というメタデータが埋め込まれており撮影日時などの情報を得ることができる。これらのメタデータを利用すれば、画像ファイルを日付ごとに分類したり、検索することは容易である。しかしながら、このようなメタ

データによる分類や検索は人間にとって必ずしも後から利用しやすいとは言えない。何故ならば人間は必ずしも日時と記録されたコンテンツの内容を明確に結びつけて覚えているわけでないからである。つまり既存の自動的に付与されるメタデータだけではファイルの内容を明確に表現するメタデータを付けることができない。問題となるのはこれらのコンテンツに対してどのようにメタデータ付けを行うのかという事である。

画像や動画にメタデータ付けを行う既存のアプローチとしては主に次の2種類がある。(1)人間が後から記録されたコンテンツに対してメタデータ付けを行う。(2)画像や動画を解析し自動的に、又はその他の要素を併用して半自動的にメタデータ付けを行う。である。(1)の例としては写真管理ソフトなどが代表的な例である。例えば Apple の iPhoto³⁾ というアプリケーションでは写真に対してキーワードを登録したり、「アルバム」という単位で写真を纏める事でメタデータ付けを行う事ができる。このようにメタデータ付けを行う事で後から簡単に検索などを行うことができる。しかしながら(1)の場合では、人間がメタデータ付けを行う際に全てのコンテンツを見て、これが一体どのような場面のものなのか人間が逐一判断する必要がある。特にファイル数が膨大な場合であったり複数の形式のファイル(動画や音声など)が含まれていると非常に

[†] 筑波大学大学院 図書館情報メディア研究科
Graduate School of Library, Information and Media Studies.

手間がかかる。(2)の例としては、松村らによる画像に人物が写っているかどうかを肌色情報を利用して抽出する手法の研究⁵⁾や北川らによる静止画像から自動的にメタデータ抽出する研究⁴⁾などが存在する。しかし、一般に画像のコンテキストを自動的に理解し抽出する事は困難である。このようなアプローチではコンテンツの分類や同期を行う際に有用な情報を取得する事は難しい。

本論文では、(1)とも(2)とも異なる、第3のアプローチを提案する。すなわち、記録するときできるだけ簡単にメタデータ付けを行うためのシステムを用意するというアプローチである。具体的には、複数のデバイスで記録された異種のメディア形式で表現されたコンテンツを「イベント」という概念に基づいてメタデータ付けを行うモデルを提案する。ここで言うイベントとは会議や打ち合せといった人間が参加する出来事の中で、本システムで各種コンテンツの分類や同期を行う括りとなるものである。本システムが扱うコンテンツの具体的な例としては、画像や映像、音声ファイルなどを対象とする。

本論文は以下のように構成される。2章では実現したい世界観の具体例を挙げる。3章ではシステムを実現するための技術的な要件について述べる。4章ではファイルの分類と同期を行うためのプロトコルについて説明する。5章ではこのシステムの応用例について述べる。

2. 実現したい世界

我々が実現したい世界では、各デバイスによる記録データに、イベントに関するメタデータ付けを簡単に行うための装置である Event ID Generator (以下 EID-Gen) を導入することがポイントである。まず、EID について説明し、次に EID を用いた世界について説明する。

2.1 Event ID Generator

図1にEID-Genのイメージ図を示す。Event ID Generatorはボタンを押すことによってイベントの開始や終了を表す信号を発行する装置である。ディスプレイとスピーカーを備えており、イベントの開始や終了をディスプレイに2次元バーコードを表示したり、スピーカーから特定の音声を流すことで表現する。2次元バーコードや音声は一つの例であり、デジタルカメラやデジタルビデオカメラなどのデバイスで記録できる物であれば、信号のエンコーディング形式は何でも構わない。

Event ID GeneratorにはIDを発行するためのボ

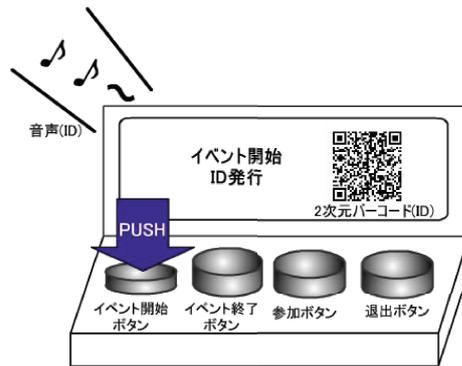


図1 Event ID Generator のイメージ図

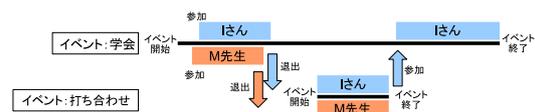


図2 シナリオ例

タンが4つある。イベント開始ボタン、イベント終了ボタン、参加ボタン、退出ボタンである。

[イベント開始ボタン] イベント開始を表す信号を発行する。

[イベント終了ボタン] イベント終了を表す信号を発行する。

[参加ボタン] イベントへの途中参加を表す信号を発行する。

[退出ボタン] イベントからの途中退出席を信号を発行する。

あるイベントを起こすためには、必ず EIG-Gen が必要である。一般には、EID-Gen は複数存在する。

2.2 具体的なシナリオ

ここでは、EID-Gen が存在する世界を、具体的なシナリオを用いて説明する。大学院生の I さんと指導教官である M 先生はある学会に参加した。I さんは自分のデジタルカメラで、M 先生はデジタルビデオカメラで会場の様子や発表を記録した。その後 I さんと M 先生は研究に関して相談する事があったので会議を途中で退出し打ち合わせを行った。打ち合わせでは、ホワイトボードにメモした内容をデジタルカメラで何枚かキャプチャした。その後 M 先生は予定があったので家に帰宅したが、I さんは学会の会場へ戻り終了時間まで発表を聞いていた。図2にこのシナリオの流れを示す。

図2のシナリオを本システムで記録するためには具体的に次のような操作を行う。なおこのシナリオで学会、I さん、M 先生が保持するものは次の通りである。

- 学会が所持するもの

- EID-Gen A
- Iさんが所持するもの
 - EID-Gen B
 - 通常のデジタルカメラ
- M先生が所持するもの
 - 通常のデジタルビデオカメラ
 - ノートPC(後述する特別なエディタが動作)

EID-Gen の利用は次のように行われる。まず、学会の主催者が、学会側にある EID-Gen A のイベント開始ボタンを押してイベントを開始する。すると、画面にイベントの開始を示す 2次元バーコードが表示される。EID-Gen A のイベント開始の 2次元バーコードは、誰もがアクセス可能な状態にしておく。例えば、会場の入り口付近に印刷したものを張ったりする。Iさんと M先生は会議に参加すると、まず、EID-Gen A の発行した 2次元バーコードを各々のデジタルカメラとデジタルビデオカメラで記録する。学会中は好きなように撮影を行う。会議を途中で退出する際には Iさんの EID-Gen B の退出ボタンを押す、画面に表示されたコードをどちらのデバイスにも記録する。その後 Iさんと M先生が打ち合わせを行う場所へ移動した際に、EID-Gen B のイベント開始ボタンを押してイベントを開始する。Iさんと M先生は学会の時と同じように画面に表示されたコードをそれぞれのデバイスに記録する。同時に、M先生は、「メタデータ作成エディタ」という特殊なエディタを PC で起動し、EID-Gen B の画面に出ているテキストで表現されたイベント開始コードを入力する。打ち合わせをしながら、M先生はそのメタデータ作成エディタに、議論のキーワードや、大事な点を記入していく。メタデータ作成エディタは、テキストが入力された際に、入力された時刻も記録している。また、それぞれのデジカメやビデオを利用して、打ち合わせを記録する。そして打ち合わせを行った後で EID-Gen B のイベント終了ボタンを押す。イベント終了の 2次元バーコードがその画面に表示されるので、それぞれのデバイスとメタデータ作成エディタに、イベント終了のコードを記録する。打ち合わせの後、M先生は帰宅するため、帰宅する。Iさんは再び学会に戻ることにしたが、EID-Gen A の参加ボタンを押す余裕がなかったため自分が持つ EID-Gen B の参加ボタンを押す、参加コードを記録する。退出する際には EID-Gen A の退出コードをデジカメに記録して、その後帰宅した。

Iさんと M先生は、自宅に帰った後、記録したコンテンツを各々が持つ PC の特別な共有ディレクトリにのこした。図 3 にその利用イメージを示す。各デバイス

の記録には撮影された写真と共に EID-Gen によって発行されたコードの記録が含まれているが、これらも含めて共有ディレクトリにのこす必要がある。そうすると、イベントごとに自動的に分類が行われ、かつ、共有ディレクトリに応じたコミュニティで共有されるのである。例えば、会議の写真を public フォルダにのこすと、その会議に参加した人々の取った写真が全て、その会議というイベントで分類され、他人からアクセス可能になる。打ち合わせの記録は他人には公開したくないため、彼らの仲間だけで共有されている group ディレクトリにのこすことにした。

数ヶ月後、M先生は I君に「えーっと、あの会議の日に打ち合わせしたよね？そのときにイベントがどうのこうのって面白い話をした気がするんだけど、なんだったっけ？」と突然言い出した。I君は、冷静に group 共有ディレクトリ中のデータに会議名と「面白い、イベント」という単語で、検索を掛けた。M先生がメタデータ作成エディタで「これは面白い」と書き込んだ時点と「イベントが重要」と書いた時点の前後に記録された画像と動画の一部がシステムによって結果として返された。I君は「この話じゃないですか？」と答えた。

3. 提案システム

2章で説明したような世界を実現するために、我々が提案するシステムのアーキテクチャは図 4 のようになる。このシステムは次のコンポーネントから構成されている。

[Event ID Generator] イベントの開始、終了、参加、退出を表現するコードを発行する。デジタルカメラなどの記録デバイスはこのコードを記録する事でイベントの情報を記録する。Event ID Generator はイベントの開始コードと終了コードを発行すると同時に、イベント ID(後述)、イベントの開始と終了時刻を Event Server に通知する。

[Event Server] Event ID Generator から全てのイベントのイベント ID、開始時刻と終了時刻の情報を受け取り記録する。

[Synchronizer] Event Server に蓄積されたイベントの情報と、共有フォルダにおかれたコンテンツ、および共に記録されたイベントに関するコードを利用する事で、コンテンツの分類を行う。

3.1 実用性に関する要件

我々は、提案システムを実用的なものとするためには、次の四つの要件を満たすことが必要であると考

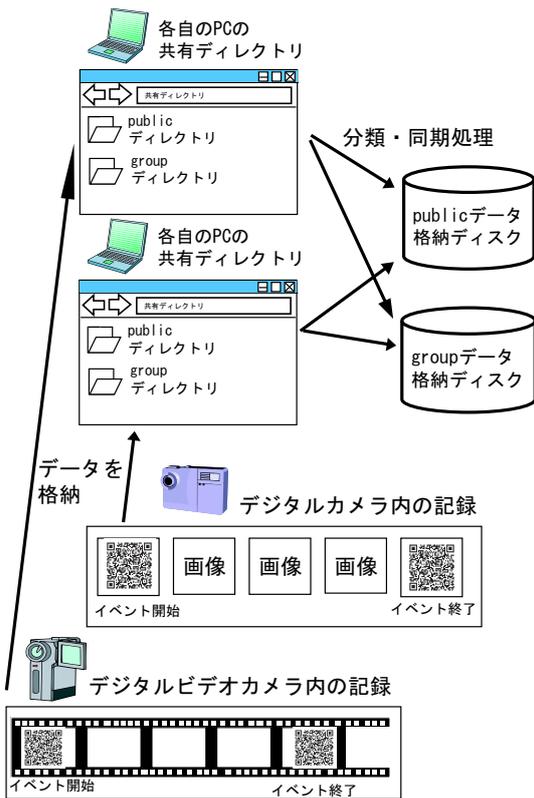


図 3 デバイス内に記録されたコンテンツの処理

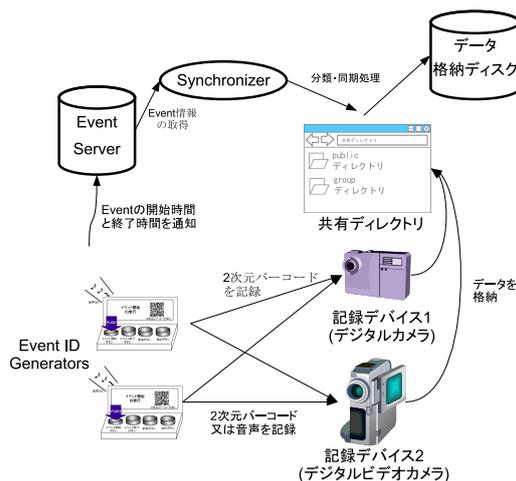


図 4 システムのアーキテクチャ

える。

- (1) 間違いをある程度許容する。

イベントの情報を記録するのは人間である。そのため押し間違い、押し忘れなどの間違いが発生する可能性がある。このような間違いをある程度修正できるような機構が必

$\langle ai \rangle_b^a \dots \dots \dots \langle /ai \rangle_b^a$

a : そのタグを発行したEvent ID Generator
 ai : Event ID Generator「a」が発行したi番目のイベント
 b : イベントを記録したデバイス

図 5 プロトコルの表現方法

要である

- (2) 複数の Event ID Generator の利用を許可する。

2.1 節で述べたようなシナリオのように、複数の EID-Gen を利用したり、複数のイベントに関わるような場合があっても期待されたとおりに振る舞わなければならない。

- (3) 利用者がイベントへの参加・退出の決定権を持つ。

利用者がイベントの途中から参加したり、途中で抜けたりという事に関しての決定権を持たなくてはならない。

4. プロトコルの提案

我々は、3.1 節で述べたような要件を満たすためには、EID Gen を用いたイベント関係コードの記録のためのプロトコルを定義することが重要だと考えた。本章では、それについて説明する。

4.1 イベントとイベントアクションシーケンス

まず、我々が扱うイベントを定義する。あるイベント e_i は 3 つ組 (EID_i, TB_i, TE_i) で表現される。ここで EID_i はイベント ID, TB_i はイベント開始時間, TE_i はイベント終了時間である。

イベントアクションシーケンスは、イベントアクションを時系列に並べたものである。ここで、イベントアクションは、イベント開始アクション、イベント終了アクション、イベント参加アクション、イベント退出アクション、匿名イベント参加アクション、匿名イベント退出アクションの 6 つのうちのいずれかである。

[イベント開始アクション] : $\langle ai \rangle_b^a$: EID-Gen a がイベント a_i を開始した。

[イベント終了アクション] : $\langle /ai \rangle_b^a$: EID-Gen a がイベント a_i を終了した。

[イベント参加アクション] : $\langle ai \rangle_b^a$: デバイス b が EID-Gen a のイベント参加コードを記録した(イベントに参加した)。

[イベント退出アクション] : $\langle /ai \rangle_b^a$: デバイス b

が EID-Gen a のイベント退出コードを記録した (イベントを退出した) .

[匿名イベント参加アクション] : $\langle \rangle_b^a$: デバイス b が EID-Gen a の匿名イベント参加コードを記録した .

[匿名イベント退出アクション] : \langle / \rangle_b^a : デバイス b が EID-Gen a の匿名イベント退出コードを記録した .

利用者による各操作と, その結果生じる EID-Gen の動作およびアクションイベントの関係を次のように決める .

- (1) EID-Gen a のイベント開始ボタンを押した時: $\langle ai \rangle_b^a$ が起こり, イベント参加コードを画面に出力する .
- (2) EID-Gen a のイベント終了ボタンを押した時: $\langle /ai \rangle_b^a$ が起こり, イベント退出コードを画面に出力する .
- (3) EID-Gen a のイベント参加ボタンを押した時: 匿名イベント参加コードを画面に出力する .
- (4) EID-Gen a のイベント終了ボタンを押した時: 匿名イベント退出コードを画面に出力する .
- (5) デバイス b が EID-Gen a のイベント参加コードを記録したとき, イベント参加アクション $\langle ai \rangle_b^a$ が起こる .
- (6) デバイス b が EID-Gen a のイベント退出コードを記録したとき, イベント退出アクション $\langle /ai \rangle_b^a$ が起こる .
- (7) デバイス b が EID-Gen a の匿名イベント参加コードを記録したとき, 匿名イベント参加アクション $\langle \rangle_b^a$ が起こる .
- (8) デバイス b が EID-Gen a の匿名イベント退出コードを記録したとき, 匿名イベント退出アクション \langle / \rangle_b^a が起こる .

デバイス b に関するイベントアクションシーケンスとは, デバイス b によって記録されたアクションだけを並べたイベントアクションシーケンスである . 図 6 に, 2.2 章のシナリオにおいての各デバイスのイベントシーケンスを示す .

デバイス b は, イベント開始アクション $\langle ai \rangle_b^a$ が生じたとき, イベント ai に参加したことになり, イベント退出アクション $\langle /ai \rangle_b^a$ が生じたとき, イベント ai から退出したことになる .

しかし, 先に説明した, 各操作とイベントアクションの関係 (7)(8) で得られた, 匿名イベント参加・退出記録アクションだけからは, どのイベントに関して参

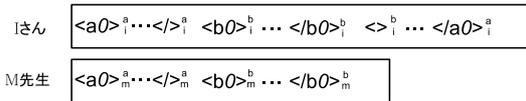


図 6 図 2 のシナリオを本プロトコルを用いて表現

加・退出されたのかわからない . また, イベント退出アクションより前にイベント終了アクションが起きているならば, 実際にはそれより前にイベントを退出していることになる .

このように, 各操作をおこなうことにより得られたイベントシーケンスから, 実際のイベントシーケンスを求める必要がある . そのために, 次から説明する制約やルールを定義する .

4.2 制 約

本プロトコルにおいて, 参加タグおよび終了タグを導出するために利用可能な制約を次のように定める .

- (1) 一つの EID-Gen が同時に起こせるイベントは唯一つである .
- (2) Event Server に問い合わせれば, 全てのイベントの開始時刻と終了時刻が分かる .
- (3) イベントに参加していないにもかかわらず, そのイベントに対する参加記録アクションが生じることはない .
- (4) イベントが終了したときには, 全てのデバイスはそのイベントから退出する .

4.3 ポリシー

上に定めた制約からだけでは解決できない (解釈が複数現れる) ものは, 次のポリシーによって解決する .

- (1) 同じイベント参加アクションが複数存在した場合, 最初のアクションを優先し, それ以外は無視する .
- (2) 同じイベント退出アクションが複数存在した場合, 最後のアクションを優先し, それ以外は無視する .

4.4 イベント参加アクションと退出アクションの導出

各デバイス b のイベントアクションシーケンス EAS_b 中のアクションを書き換え, イベント参加アクションと退出アクションを導出する . その手順は次の通りである .

- (1) Event Server にアクセスし, EAS_b にイベント開始アクションとイベント終了アクションを追加する .
- (2) イベント退出アクションよりも先に同じイベントの終了アクションがあった場合, イベント終了アクションの直前にイベント退出アクション

- を移動する。
- (3) 次の規則により、匿名イベント参加アクションをイベント参加アクションに変換する。
- ある匿名イベント参加アクション $\langle \rangle_b^a$ より前に、終了していないイベント開始アクション $\langle ai \rangle_a^a$ が存在するならば、 $\langle \rangle_b^a$ をイベント参加アクション $\langle ai \rangle_b^a$ に変換する。
 - 匿名イベント参加アクション $\langle \rangle_b^a$ よりまえに、終了していないイベント開始アクション $\langle ai \rangle_a^a$ が存在しなければ、 $\langle \rangle_b^a$ より後でかつ最も近いような、イベント参加アクション $\langle ci \rangle_b^c$ で $\langle \rangle_b^a$ を置き換える。
- (4) 次の規則により、匿名イベント退出アクションをイベント退出アクションに変換する。
- もし、ある匿名イベント退出アクション $\langle \rangle_b^a$ より前に、終了していないイベント開始アクション $\langle ai \rangle_a^a$ が存在するならば、 \langle / \rangle_b^a をイベント退出アクション $\langle /ai \rangle_b^a$ に変換する。
 - 匿名イベント退出アクション \langle / \rangle_b^a よりまえに、終了していないイベント開始アクション $\langle ai \rangle_a^a$ が存在しなければ、 \langle / \rangle_b^a を、その時点でデバイス b が参加している全てのイベントに関するイベント終了アクションの列で置き換える。

5. シナリオ例におけるシステムの動作

本章では2章で説明したシナリオ例が4章の枠組みを用いるとどのように処理されるのかを説明する。ここでIさんが所持するデジカメを i で表し、M先生が所持するデジタルビデオカメラを m で表す。

- (1) EID-Gen A のイベント開始ボタンを押し、「学会」イベントがスタートした時
- $\langle a0 \rangle_a^a$ が起こりイベント参加コードが画面に出力される。
- (2) IさんとM先生が各々のデバイスでEID-Gen A が発行した参加コードを記録した時
- IさんとM先生各々のデバイス内にイベント参加アクション $\langle a0 \rangle_i^a$ と $\langle a0 \rangle_m^a$ が記録される。
- (3) IさんとM先生が退出ボタンを推して「学会」イベントを退出する時
- IさんとM先生各々のデバイス内に匿名イベント退出アクション \langle / \rangle_i^a と \langle / \rangle_m^a が記録される。

- (4) EID-Gen B のイベント開始ボタンを押し、「打ち合わせ」イベントが開始された時
- $\langle b0 \rangle_b^b$ が起こり、イベント参加コードが画面に出力される。
- (5) IさんとM先生が各々のデバイスでEID-Gen B が発行した参加コードを記録した時
- IさんとM先生各々のデバイス内にイベント参加アクション $\langle b0 \rangle_i^b$ と $\langle b0 \rangle_m^a$ が記録される。
- (6) EID-Gen B のイベント終了ボタンを押し、「打ち合わせイベント」が終了した時
- $\langle /b0 \rangle_b^b$ が起こり、イベント退出コードが画面に表示される。
- (7) IさんとM先生が各々のデバイスで退出コードを記録した時
- IさんとM先生各々のデバイス内にイベント退出アクション $\langle /b0 \rangle_i^b$ と $\langle /b0 \rangle_m^a$ が記録される。
- (8) IさんがEID-Gen B の参加ボタンを押した時
- 匿名イベント参加コードが画面に表示される
- (9) Iさんがデバイスで匿名イベント参加コードを記録する時
- Iさんのデバイスに $\langle \rangle_i^b$ が記録される。
- (10) EID-Gen A のイベント終了ボタンが押された時
- $\langle /a0 \rangle_a^a$ が起こりEID-Gen A の画面にイベント退出コードが表示される。
- (11) IさんがデバイスでEID-Gen A が発行したイベント退出コードを記録した時
- イベント退出アクション $\langle /a0 \rangle_i^a$ がIさんのデバイスに記録される。

このシナリオにおいてイベント参加アクションおよびイベント退出アクションを導出しなければならないのは、IさんとM先生が学会イベントから退出した際の匿名イベント退出アクション \langle / \rangle_i^a , \langle / \rangle_m^a と、Iさんが学会に参加した際の匿名イベント参加アクション $\langle \rangle_i^b$ である。

匿名イベント退出タグ \langle / \rangle_i^a , \langle / \rangle_m^a は4.4節の(4)の規則を用いる事でそれぞれイベント退出アクション $\langle /a \rangle_i^a$, \langle / \rangle_m^a となる。匿名イベント参加アクション $\langle \rangle_i^b$ は4.4節の(3)の規則を用いることでイベント参加アクション $\langle a0 \rangle_i^b$ となる。このようにしてイベント参加アクション及びイベント退出アクションが導出された結果を図7に示す。

1さん `<a0>a...</a0>a <b0>b ... </b0>b <a0>b ... </a0>a`
 M先生 `<a0>a...</a0>a <b0>b... </b0>b`

図 7 イベント参加アクション及びイベント退出アクションを導出した結果

6. おわりに

本稿では、複数のデバイスで記録されたコンテンツの分類と同期を行うためのシステムについて提案した。ポイントは、特別な仕組みを持たないような各種デバイスによって記録されたコンテンツであっても、容易に分類・同期を行うための仕組みを用意したことである。具体的には、Event ID Generator という装置を用いて、各種デバイスで記録されたデータに、参加したイベントに関するメタデータを追加する。本稿では、実現のためのアーキテクチャ、Event ID Generator を利用するためのそしてプロトコルについて提案した。今後はプロトタイプシステムの構築などを行う予定である。

謝 辞

ゼミなどでご議論いただきました筑波大学大学院図書館情報メディア研究科の田畑孝一教授、阪口哲男助教授、永森光晴講師に感謝致します。本研究の一部は日本学術振興会科学研究費補助金若手研究 (B)(課題番号 15700108) による。

参 考 文 献

- 1) JEITA CP-3451, Exchangeable image file format for digital still cameras: Exif Version 2.2, <http://tsc.jeita.or.jp/avs/data/ep3451.pdf>
- 2) Michael J. Franklin: Challenges in Ubiquitous Data Management. Informatics 2001: 24-33
- 3) iPhoto, <http://www.apple.com/jp/ilife/iphoto/>
- 4) 北川高嗣, 中西崇文, 清木康: “静止画像メディアデータを対象としたメタデータ自動抽出方式の実現とその意味的画像検索への適用,” 情報処理学会論文誌: データベース, VOL.43, No. SIG12(TOD16), pp38-51, 2002.
- 5) 松村, 岩井, 谷内田: “肌色情報を用いた複数人物追跡”, 情報処理学会 CVIM 研究会報告, NO.133. pp. 133-138(2002).