

異世界放送システムのための映像処理ルール記述方式

松本 哲^{†1} 義久智樹^{†1} 川上朋也^{†2,†1} 石芳正^{†3,†1} 寺西裕一^{†4,†1}

近年、USTREAM やツイキャスといった、個人がインターネットを介してリアルタイムな映像配信を行えるインターネットライブ放送サービスが普及している。我々の研究グループでは、異世界放送システムと呼ぶ、分散型映像処理を用いたインターネットライブ放送システムを研究開発している。異世界放送システムでは、計算能力の高い映像処理サーバを用いることで、映像処理にかかる時間を短縮しつつ、映像データを配信できる。インターネットライブ放送では、放送禁止区域に入ると画面を黒くする、放送者以外が写っていればモザイクをかける、といった放送者の意図に応じた映像処理を行うことがある。これらの映像処理を自動的に行うことで放送者の操作回数を削減できるが、これまでの異世界放送システムでは、放送開始前に設定された映像処理を放送者や視聴者の指示に従って付加することしかできなかった。そこで本研究では、放送者の意図に応じて柔軟に自動的に映像処理を行える異世界放送システムのための映像処理ルール記述方式を提案する。提案方式では、映像処理を行うタイミングと条件をルール形式で記述することで、柔軟かつ自動的な映像処理の実現を目指す。

1. はじめに

近年の映像配信技術の発達に伴い、USTREAM やツイキャスといった、個人がインターネットを介してリアルタイムな映像配信を行えるインターネットライブ放送サービスが普及している[1]。例えば、USTREAM では、個人のパソコンで USTREAM のホームページにブラウザでアクセスし、放送を開始するボタンをクリックすることで、パソコンに接続されたカメラで撮影された映像を配信できる。ツイキャスでは、主にスマートフォンからの配信を対象としており、専用アプリを用いてスマートフォンのカメラで撮影された映像を配信できる。我々の研究グループでは、異世界放送システムと呼ぶ、ライブ映像に対し、加工処理等をリアルタイムに行なうことで現実のライブ放送ではなく、あたかも異なる世界のライブ放送であるかのような効果を生み出すことを可能とする分散型映像処理を用いたインターネットライブ放送システムの研究をしている(図1)[2]。異世界放送システムでは、インターネットライブ放送を行う個人(放送者)の撮影端末で取得した映像データを、映像処理サーバ(異世界放送サーバ)に送信する。異世界放送サーバは、受信した映像データに対して映像処理を実行して放送サーバに送信し、放送サーバが視聴者に映像データを配信する。計算能力の高い異世界放送サーバを用いることで、映像処理にかかる時間を短縮しつつ、映像を配信できる。インターネットライブ放送では、放送禁止区域に入ると画面を黒くする、放送者以外が写っていればモザイクをかける、といった放送者の意図に応じた映像処理を行うことがある。映像処理を柔軟に変更することで、放送者のもつ様々な意図を反映した映像処理を行える。また、これらの映像処理を自動的に行うことで放送者の操作回数を

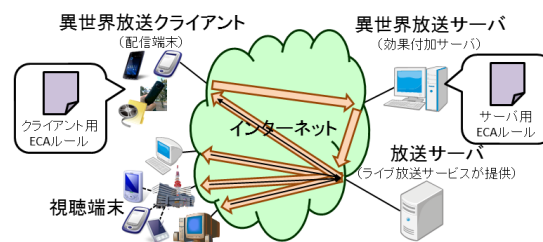


図1 異世界放送システムのシステム構成
Figure 1 The system architecture of other world broadcasting system

削減でき、放送時の放送者の身体的負担を軽減できる。しかし、これまでの異世界放送システムでは、放送前に設定された映像処理を放送者や視聴者の指示に従って付加することしかできなかった。

そこで本研究では、放送者の意図に応じて柔軟に自動的に映像処理を行える異世界放送システムのための映像処理ルール記述方式を提案する。提案方式では、映像処理を行うタイミングと条件およびその処理を ECA ルールで記述する。ECA ルールとは、システム内で事象 (Event) が起こり、ある特定の状態 (Condition) を満たしているとき、その行動 (Action) を実行する、という動作原理に基づくルールである。提案方式では、ECA ルールを用いることで、柔軟かつ自動的な映像処理を実現する。

以下、2章で関連研究について説明し、3章で本研究で提案する映像処理ルール記述方式の設計と実装について述べ、4章で本稿をまとめる。

2. 関連研究

SMIL(Synchronized Multimedia Integration Language)3.0では、映像を提示するタイミングや付加する映像効果等をXML (eXtensible Markup Language)で記述する[3]。例えば、映像に他の絵を重畳表示させたり、テロップを出すといった映像処理を記述できる。しかし、SMILで規定された映像処理しか行えない。異世界放送システムでは、サービス提供者や放送者が実装したプログラムライブラリ(異世界放送ライブラリ)を用いることで様々な映像処理を行える

^{†1} 大阪大学

Osaka University

^{†2} 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

^{†3} 株式会社 PIAX

PIAX Inc.

^{†4} 情報通信研究機構

National Institute of Information and Communications Technology

ため、これらを実行できる映像処理記述方式が求められる。

TVML は、3次元映像を生成する言語であり、出演キャラクターや物体の動作等を記述できる[4]。ある映像に対する処理の付加を対象としておらず、インターネットライブ放送の映像処理記述方式として利用できない。

文献[5]では、小型装置の制御に ECA ルールを用いている。簡易な ECA ルールを複数組み合わせることで複雑な制御を可能にしている。小型装置の制御用に設計された ECA ルールであり、映像処理に関するルールを記述できない。

3. 映像処理記述言語の設計と実装

本章では、まず初めに筆者らの研究グループで研究開発してきた異世界放送システムについて説明し、その後映像処理記述方式の設計と実装について述べる。

3.1 異世界放送システム

異世界放送システムのシステム構成を図 1 に示す。異世界放送クライアントは、インターネットライブ放送の配信者が使用するスマートフォンやタブレットといった配信端末であり、接続されているカメラやマイクから映像と音声を取得する。異世界放送サーバは映像効果や音声効果を付加するサーバである。異世界放送サーバは、配信端末から映像データや音声データを受信して効果付加処理を行い、インターネットライブ放送サービスが提供する放送サーバに送信する。効果に関するデータも受信する。放送サーバは、受信している配信端末のインターネットライブ放送を視聴している視聴端末にデータを配信する。視聴端末は、インターネットライブ放送サービスが提供している配信一覧の中から興味のある配信を選択して視聴する。

近年のクラウドコンピューティング環境の普及[6]に伴い、インターネットライブ放送サービス提供者が高性能な計算機を設置し、放送者が利用する上記のようなシステム構成は現実的な構成といえる。

3.2 要求機能

本研究では、以下を異世界放送システムのための映像処理記述方式への機能要求と考える。

- 映像処理の柔軟な変更：インターネットライブ放送では、放送禁止区域に入ると画面を黒くする、放送者以外が写っていればモザイクをかける、といった放送者の意図に応じた映像処理を行うことがある。映像処理の内容は放送者によって異なる場合があり、例えば他の放送者では、放送禁止区域に入ると放送自体を停止したり、許可を得ているために放送者以外が写ってもそのまま放送することがある。映像処理を柔軟に変更することで、放送者のもつ様々な意図を反映した映像処理を行える。
- 自動的な映像処理：放送者や視聴者の指示で付加する映像処理に加えて、自動的に映像処理を行うことで、放送者の操作回数を削減でき、放送時の放送者の身体

表 1 イベントの種類

イベント名	内容
Get_Frame	映像の 1 フレームの画像データを取得した
Receive_Msg	メッセージを受信した
Get_Data	センサデータを取得した
Timer_Expire	設定したタイマーがタイムアウトした

表 2 コンディションの種類

コンディション変数	内容
Num_Find_Object	取得した画像データで認識されたオブジェクトの数。認識するオブジェクトは、Object パラメタで指定する。
Brightness	取得した画像データの平均明度
Contrast	取得した画像データの平均コントラスト
Msg_Source	メッセージの送信元
Msg_No	メッセージの番号
Data_Type	取得したデータのタイプ
Data_Value	取得したデータの値
Timer_No	タイムアウトしたタイマーの番号

表 3 アクションの種類

アクション名	内容	パラメタ
Surround_Object	認識したオブジェクトを囲む矩形を重畳表示する。	Surround_Color : 色 Thickness : 太さ
Blur_Object	認識したオブジェクトを囲む矩形の範囲をぼかす。	Blur_Value : ぼかしの強さ
Change_Brightness	取得した画像データの平均明度を変更する。	Value:設定する平均明度
Change_Contrast	取得した画像データの平均コントラストを変更する。	Value:設定する平均コントラスト
Start_Broadcasting	放送を開始する。	なし
Stop_Broadcasting	放送を終了する。	なし
Send_Msg	メッセージを送信する。	Destination:送信先 Msg_No:メッセージ番号
Set_Timer	タイマーを設定する。	Timer:タイマーの番号

的負担を軽減できる。例えば、画面を黒くする、放送者以外が写っていればモザイクをかける映像処理が自動的に行われる。

3.3 ECA ルールによる動作記述の設計

3.2 節で述べた要求事項を満たす映像処理記述の動作原理として、ECA ルールが挙げられる。1 章に述べた通り、ECA ルールでは、イベントが発生し、コンディションが満たされている場合にアクションが実行される。簡素な ECA ルールを組み合わせることで複雑な映像処理を実現できる。放送者の意図に応じて ECA ルールを組み合わせることで、柔軟に映像処理を変更でき、初めの要求事項を満たせる。また、イベントの発生を異世界放送クライアントや異世界放送サーバが認識することで自動的に映像処理を行え、2 番目の

要求事項を満たせる。そこで、本研究では、異世界放送システムにおける映像処理の記述に ECA ルールを用いる。

3.4 ECA ルールを用いた映像処理記述の設計と実装

本節では、異世界放送システムにおける ECA ルールを用いた映像処理記述の設計と実装について詳述する。ECA ルールは、異世界放送クライアントや異世界放送サーバに保存する。

3.4.1 イベント (Event)

イベントに記述できる内容を表 1 に示す。メッセージとは、異世界放送クライアントや異世界放送サーバが、同じネットワークに接続されている他の端末から送信される通知用のデータであり、メッセージ番号を介して他の端末への映像処理の指示等を行える。Receive_Msg にて指示のイベントを発生させる。また、これらの端末に接続されている温度センサや位置センサといったセンサがデータを取得した際に発生するイベントの Get_Data もある。タイマーも利用でき、タイマーのタイムアウトをイベントに指定できる。

3.4.2 コンディション (Condition)

コンディションに記述できる内容を表 2 に示す。コンディション変数とは、コンディションの記述部分で参照できる変数であり、比較演算を行える。例えば、Num_Find_Object 変数は Object 引数を持ち、Object が顔の検出器を示しているれば、Num_Find_Object 変数は取得した画像データに写っている顔の数を示す。これらのコンディションを複数組み合わせ合わせて利用できる。

3.4.3 アクション (Action)

アクションに記述できる内容を表 3 に示す。オブジェクトを囲む矩形を重畳表示したり、ぼかすといった映像処理の他、放送を開始や終了といったインターネットライブ放送に関する動作を記述できる。アクションの詳細部分についてはパラメタで記述できる。コンディションと同じく、複数のアクションを組み合わせて利用できる。

3.4.4 ECA ルールの処理方法

異世界放送サーバや異世界放送クライアントは、保持しているルールに記述されたイベントの発生を監視する。イベントが発生すると、コンディションが満たされているか判断する。コンディションが満たされている場合には、アクションを実行する。本研究では、ECA ルールの記述に JSON (JavaScript Object Notation) 形式を用いた。JSON 形式は、JavaScript にて利用できる記述方式であり、多くの Web アプリケーションで利用されている。ECA ルールを JSON 形式で記述することで、将来的に他の Web アプリケーションに応用しやすくなる。一例を図 2 に示す。この例では、“Get_Frame” が発生し、コンディションの評価を行う。コンディションの object パラメタには人の顔を指す分類器が指定されており、“Num_Find_Object” で画像データに含まれる人の顔の数を参照できる。この値が 1 以上であれば、顔にぼかしをかけるアクションを実行する。

```
{
  "ルール A": {
    "eventName": "Get_Frame",
    "condition": {
      "name": "Num_Find_Object",
      "object": "http://object1_haar.xml",
      "value": ">=1"
    },
    "action": {
      "name": "Blur_Object",
      "Blur_Value": "60"
    }
  },
  "ルール B": {
    "eventName": "Get_Frame",
    "condition": {
      "name": "Brightness",
      "param": "30"
    },
    "action": {
      "name": "Change_Brightness",
      "value": "128"
    }
  }
}
```

図 2 JSON 形式映像処理ルールの記述例

Figure 2 An example of video processing descriptions by JSON rule format

“object” の周辺映像にぼかし効果を “Blur_Value” の値でアクションを実行する。これを “ルール A” と名付けている。ルール B は明るさを変更するルールである。

4. おわりに

本研究では、異世界放送システムのための映像処理ルール記述方式の設計と実装を行った。提案方式では、映像処理を行うタイミングと条件を ECA ルールを用いて記述する。ECA ルールを用いることで、放送者の意図に応じて柔軟かつ自動的な映像処理を実現する。

今後、ECA ルールの判定時間や映像処理時間について、ECA ルールを用いない場合との比較評価や、他の映像処理記述方式との機能的な比較評価を行う。また、受信側の意図を映像に反映できる仕組みも考慮する必要がある。

謝辞

本研究の一部は、文部科学省科学研究費補助金・基盤研究(B)(15H02702)、挑戦的萌芽研究(26540045)の研究助成によるものである。ここに記して謝意を表す。

参考文献

- [1] Sodagar, I., “The MPEG-DASH Standard for Multimedia Streaming Over the Internet”, IEEE Multimedia, Vol. 18, No. 4, pp. 62-67 (2011).
- [2] 義久智樹, 川上朋也, 石芳正, 寺西 裕一, “異世界放送: 映像効果と音声効果を伴う分散型インターネットライブ放送システム”, 情報処理学会シンポジウムシリーズ マルチメディア, 分散, 協調とモバイルシンポジウム(DICOMO2016)論文集, pp. 1827-1832 (2016).
- [3] Synchronized Multimedia Integration Language (SMIL 3.0) W3C Recommendation 01 December 2008, <http://www.w3.org/TR/SMIL/> (2008).
- [4] Kidawara, Y., Yamamiya, T., Hayashi, M., and Tanaka, K., “Fourth International Conference on Creating”, IEEE Connecting and Collaborating through Computing (C5'06), pp.2-9 (2006)
- [5] J.Widom and S.Ceri: Active Database Systems, “Triggers and Rules for Advanced Database Processing”, Morgan Kaufmann Publishers, (1996).
- [6] Cheng, B., “MediaPaaS: A Cloud-Based Media Processing Platform for Elastic Live Broadcasting”, Proc. of 2014 IEEE International Conference on Cloud Computing (CLOUD), pp. 713-720 (2014).