

# 診療映像における自動イベント検出の検討

浅井 博次 田中 等幸

## Investigation of Automatic Event Detection in Intervention Video

Hirotsugu ASAI Tomoyuki TANAKA

**あらまし** 動画による診療記録は事実をありのまま記録することができるため、医療行為の透明性確保や信頼性、医療技術の向上に有効である。しかし、動画の内容確認や見たい箇所の検索に時間がかかり、気軽に活用することが困難である。撮影されている映像内容などのメタデータを動画データに付加し、メタデータを基にデータ検索を行う手法が多数検討されているが、メタデータ入力の手間が問題となっている。本研究では、イベント（医療行為）を一つのメタデータとして自動的に付加することを目指し、イベント検出手法の検討を行う。具体的には映像情報を基にした手法と超音波三次元タグを用いた三次元位置情報に基づく手法の二つのイベント検出手法について検討したので報告する。

**キーワード** 救急医療、診療記録、動画、イベント検出、三次元位置情報

### 1. はじめに

本研究では、動画を用いた診療記録により、従来に比べより詳細な記録を残すことで医療の透明性の確保や医療技術の向上に寄与することを目指している。

一般的な診療記録であるカルテでは、実施した行為や所見の要約がテキストベースで記録されており、短時間で患者情報を把握することが可能である。しかし、記録内容が記録者の主観に依存してしまうため、記録の量・質・信頼性を確保することが困難であるという問題点がある。医療の現場で実際に何がどのように行われたのかわからないため、下された判断・実施された処置が妥当であったかどうかをカルテ情報だけでは判断できない。

特に、手術や救急処置など重要な医療行為では、患者の容態、処置のタイミング、処置のやり方などが複雑に影響するため、事後検証や処置データベースなどへの活用には客観的かつ詳細な事実の記録が必要となる。動画を用いた診療記録では、診療の様子をありのまま記録することが可能であるため、これらの要求を満たす有効な手段であると言える。しかしながら、動画記録の運用には主に2つの課題が存在する。1つは、必要な部分の映像を的確に撮影することである。撮影すべき対象は時々刻々と変化し、時に医療器具や医師、看護師等により遮蔽されてしまうため、有効な映像を常に記録することが困難である。もう1つは、動画の内容確認に時間がかかり、手軽な活用ができないことである。前者に対しては、複数カメラの使用により多様なアングルから対象を撮影することが考えられるが、多数のカメラを状況に応じて適切に制御するのは容易なことではない。後者に対して

は、処置イベントなどのメタデータを豊富に付加し、内容要約の作成や、見たい場所へのアクセスを容易にするための補助機能などを提供することで容易な動画記録活用を可能にすることが重要である。しかし、メタデータを付加する作業は映像内容確認作業以上に手間のかかる作業であり、実行することが困難である。そこで我々は、標準的な治療手順などを規定した標準的治療ガイドラインを活用したカメラ制御・処置内容記述機能により、カメラ制御・メタデータ付加双方における作業コスト（手間）を抑えたプロトタイプシステムの開発を進めてきた。昨年度までは、人による操作を前提として作業コスト削減を検討した<sup>[1]</sup>。本年度は映像やセンサ情報から処置などのイベントを検出しメタデータとして自動付加することによる更なる作業コスト削減を目指し、映像情報を基にしたイベント検出手法と超音波三次元タグを用いた三次元位置情報に基づくイベント検出手法について検討したので報告する。

### 2. 映像によるイベント検出

#### 2.1 はじめに

診療の様子を動画で記録することが前提であることから、まず、映像情報からの処置イベント検出を検討した。

一般的に何らかのイベントを検出するためには、誰が、何（何処）に対して、どの道具を用いて、どんな行為（動き）をするか、を理解する必要がある。つまり、人や医療機器など特定の対象を識別し、その位置や動きによって処置を定義しなければならない。しかし、映像中の対象サイズの変化や光源変動により、対象を精度良く識別

することは容易ではない。実利用を想定した場合、安定したイベント検出性能が重要であることから、対象識別を必要としない手法が望ましい。一方、医療行為では、患者のどの部分に対する処置であるかが重要な意味を持つことから、特定の場所における動き情報の識別を処置イベント検出に活用できると思われる。そこで、本研究では、対象識別を必要とせず、特定領域の動き情報のみを用いたイベント検出手法を胸骨圧迫心臓マッサージ（以下、心臓マッサージ）に適用し検討した<sup>[1]</sup>。本手法は対象識別を必要としないため、安定したイベント検出が期待できる。

## 2. 2 心臓マッサージ

心臓マッサージは救命措置における心肺蘇生法（CPR）の中で実施される処置であり、心停止状態の患者に対し、患者胸部側面に立ち、一定のリズムで継続的に患者の胸部を押す・休止という単純反復動作を行うものである。1分間に100回のペースで心臓マッサージを実施すること、心臓マッサージと換気を30：2の割合で実施すること、CPR中の心臓マッサージ中断時間をできる限り少なくすること、などとガイドライン<sup>[2]</sup>により詳細に実施方法が規定されており、これらの規定情報を心臓マッサージ検出に利用することができる。更に、心臓マッサージのように周期性のある特徴的な動きを伴う処置が他に見られないことから、処置映像から抽出した動き情報の周波数解析を行うことで心臓マッサージを実施している時間領域（以下、心臓マッサージ実施区間）を検出する。

胸部を押すという処置内容から、多くの処置は心臓マッサージを実施していない時にしか行うことができない。そのため、標準的治療ガイドラインと関連付けることにより、心臓マッサージ実施区間を検出するだけでCPRの処置の流れをおおまかに記述した処置映像要約が可能となる。心臓マッサージ実施区間の検出は、心臓マッサージ中断区間情報を基にしたCPRの評価にも活用することが可能である。

## 2. 3 検出方法

図1に検出フローを示す。

始めに、患者の位置を基準として心臓マッサージを実施している画像領域を設定する。本研究では手入力で領域指定を行った。

次に、指定した領域の動き情報を抽出する。動き情報

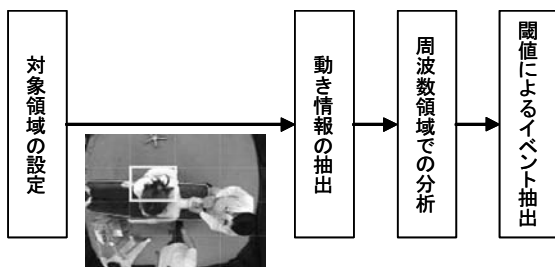


図1 検出フロー

の抽出には、指定した映像領域内における前後フレーム間の輝度に関するSAD(Sum of Absolute Difference)を採用した。図2に示すように、心臓マッサージ実施区間ではSADの値が大きく変化しており、非実施区間とは明確な違いが見られる。また、実施区間では顕著な周期性が見て取れる。

周期性を利用し、本研究ではSADデータに対し短時間フーリエ変換（STFT）を実行し、動き情報の周波数解析を行うことで、精度良く実施区間の検出を行った。STFTは一定の大きさの窓関数を少しずつずらしながら信号を切り出し、切り出した信号をフーリエ変換してスペクトルを計算する手法である。得られるスペクトル列から周波数領域での信号の時間的変化を解析することが可能である。

図3に横軸を分単位の周波数に変換した心臓マッサージ実施区間と非実施区間のスペクトル例を示す。図3(a)に示す通り、心臓マッサージ実施区間のスペクトルには、ガイドラインに規定してある100回/分付近に鋭いピーク

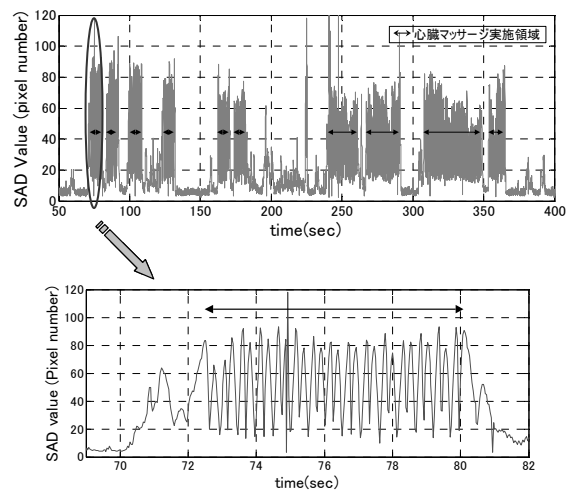
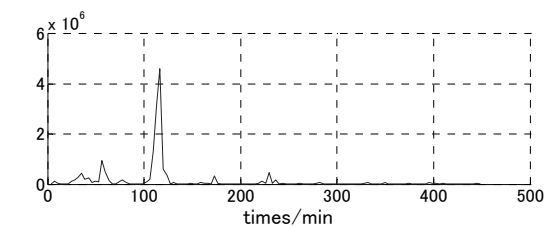
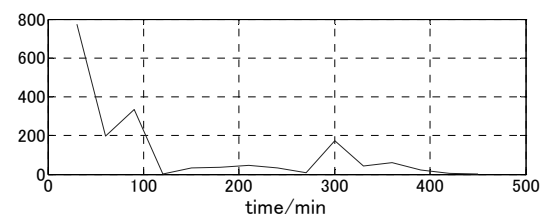


図2 SADによる動き情報の時間推移



(a) 心臓マッサージ実施区間



(b) 心臓マッサージ非実施区間

図3 スペクトル例

が見られ、心臓マッサージによる運動の周期性がよく抽出できていることがわかる。一方、非実施区間のスペクトルは全体的に明確なピーク値が見られず、ピークの値も非常に小さい。検出対象である心臓マッサージ以外に同様の反復動作を伴う処置がCPRにないことから、スペクトルのピーク周波数値を基に、ガイドラインに規定される周波数付近で一定時間以上継続してピークが見られる区間を心臓マッサージ実施区間として抽出する。

### 2. 3 実験

シミュレーション人形を用いた心肺蘇生講習会において、処置を真上から撮影した一次救命処置 (BSL) の映像を用いて本手法の有効性を検証した。使用した映像は160×120ピクセル、30fpsの非圧縮AVIファイルである。同じ位置から撮影した映像でも処置者の身長・体格などの身体的特徴により画像における処置者の占める領域が大きく異なる。また、処置者の技術レベルの差などから処置時の動きも異なる。これらの影響を考慮し、本研究では患者であるシミュレーション人形の位置を基に対象とする映像領域のサイズを大小2種類設定した。また、STFTの窓関数に矩形窓を採用し、異なる3種類の窓サイズを用いて実験を行い、窓サイズの影響を検討した。図4(a)に対象領域を、図4(b)と図4(c)に異なる処置者の画像例を示す。図4(a)の黒枠及び白枠領域が設定した対象領域である。

表1、表2に実験結果を示す。表1では、実際に行われた心臓マッサージ実施数のうちの提案手法による検出数を示す。表2では、実際に心臓マッサージが行われた時間を積算し、正しく検出された時間の割合を示す。DATA1~3はそれぞれ異なる処置者による処置映像である。DATA1は小柄な処置者、DATA3は背の高い処置者である。表1から対象領域サイズによって検出数が大きく異なっていることが分かる。対象領域サイズが大きいDATA1のケースでは対象領域に占める処置者領域の比率が低いことから、処置者以外の領域の動きの影響を受けるのが原因と思われる。一方、対象領域サイズが小さいDATA3のケースでは背中部分など処置者の一部しか対象領域に含まれていないため、処置者の動き情報を十分に抽出できていないことが原因と考えられる。対象領域のサイズの違いによる検出精度の低下は表2では更に顕著に現れており、処置者に合わせて対象領域を適切に設定することが非常に重要であることがわかった。

窓サイズによる影響は検出数にはあまり現れていないが、検出時間には特定の傾向が見られる。対象領域設定



図4 対象領域、及び画像例

の影響を排除し、窓サイズの検出性能への影響について検討するため、それぞれのデータについて検出精度が良い方の対象領域の検出結果に着目すると、窓サイズの拡大に伴い検出率が低下している。これは、窓サイズの拡大に伴うイベント検出の時間分解能低下に起因していると考えられる。また、対象とするデータの時間領域が広がることで、1つのスペクトルの中に異なる複数の動き情報が入り込む可能性が高まり、心臓マッサージの動きを抽出することが困難になることも考えられる。計算コストと検出精度を考慮すると、心臓マッサージの動きを抽出できるだけの小さい窓サイズが最もよい。今回の実験では、最小設定の1.06秒の窓サイズでも精度良く心臓マッサージイベントの検出ができることが確認できた。

検出時間による評価が低いDATA3について、心臓マッサージイベント毎に検出結果と実際の実施区間とを詳細に比較した結果、最も精度の良い設定(窓サイズ1.06sec, 対象領域:大)で、実際に実施した時間と検出した心臓マッサージイベントの開始・終了時間との間に1秒前後のズレがある場合が多数あり、その積み上げが検出時間による検出率を押し下げていることが確認できた。ガイドラインでは心臓マッサージを1分間に100回という非常に早いリズムで実施するように規定しているが、当該の処置者は、開始・終了時にリズムが緩くなる傾向が見られ、これが検出時間における1秒前後のズレの原因と思われる。本実験では、心臓マッサージのための動作をはじめた時間を開始時間、動きがとまった時間

表1 検出結果 (検出数による評価)

	STFT 窓サイズ (sec)	対象映像領域サイズ			
		小		大	
		検出数	誤検出数	検出数	誤検出数
DATA1	1.06	10/10	0	6/10	3
	2.13	10/10	0	6/10	1
	4.27	10/10	0	4/10	0
DATA2	1.06	9/11	0	10/11	0
	2.13	10/11	0	10/11	0
	4.27	10/11	0	10/11	0
DATA3	1.06	13/33	0	30/33	0
	2.13	27/33	0	29/33	0
	4.27	23/33	0	27/33	0

表2 検出結果 (検出時間による評価)

	STFT 窓サイズ (sec)	対象映像領域サイズ			
		小		大	
		検出率	誤検出率	検出率	誤検出率
DATA1	1.06	89.84%	0.18%	35.11%	9.78%
	2.13	87.78%	0.67%	38.46%	4.20%
	4.27	78.45%	4.65%	34.38%	0.47%
DATA2	1.06	74.67%	0.38%	92.75%	0.38%
	2.13	87.97%	0.27%	86.53%	0.27%
	4.27	80.65%	2.35%	81.65%	3.16%
DATA3	1.06	28.44%	0.13%	72.33%	0.61%
	2.13	60.60%	1.67%	66.22%	2.47%
	4.27	46.50%	8.14%	52.00%	7.64%

を終了時間と規定しているため、このズレに妥当な結果であると言える。

本実験により、対象領域サイズ、および STFT の窓サイズの検出性能への影響を確認することができた。対象領域の設定は検出性能への影響が大きく、処置者に応じた適切な設定が重要であることがわかった。本実験では手動で対象領域の設定を行っているが、画像認識技術などの活用により自動化することで安定した検出性能を実現できると思われる。また、STFT の適切な窓サイズ (1.06sec) が判明した。今回の実験について、精度のよい方の対象領域サイズ、且つ適切な窓サイズである 1.06sec の実験結果のみに着目すると、検出件数による評価では、全てのデータにおいて約 90% の精度を、検出時間による評価でも 70% 以上の検出性能が得られており、提案手法により精度良く心臓マッサージイベントを検出することが可能であることが確認できた。今後、画像認識による対象領域設定の自動化でどの程度適切な設定が可能であるかを検証していく予定である。

### 3. 位置情報によるイベント検出

#### 3. 1 位置情報活用の必要性

2章では映像から動き情報を分析しイベントを検出する手法を示した。映像を基にイベントを検出する手法の問題点として、照明環境の変動、適切な撮影範囲の決定、人や物による遮蔽などが原因となり、イベント検出が困難な場合がある。そこで、本章では映像を用いず、上述した問題を克服するためにセンサ情報を利用し、イベント検出を行う手法について述べる。具体的には、三次元の位置情報を取得できるセンサを医療機器や処置者に装着し、取得された時系列の位置情報とそれぞれの振る舞いを関連付けることでイベントを検出する。

#### 3. 2 超音波三次元位置測定システム

映像以外に位置情報を得る方式としては、電波、赤外

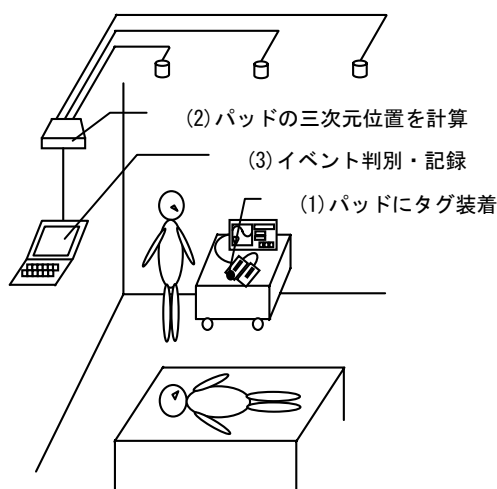


図5 検出フロー

線、磁気などの電磁波を利用して、送受信器で三次元位置情報を取得する方法がある。その中で、電波や赤外線による方式ではゾーン検出が主な利用目的として使用され、その検出精度はおよそメートル単位のオーダーである。磁気方式はミリ単位オーダーの高い検出精度である一方で、付近に金属等の磁性体がある場合に、その影響によって安定した三次元位置情報を取得できない問題がある。我々が想定する環境は金属等の磁性体が多く存在し、それらの物体を排除することが不可能な救急治療室である。そこで、安価でかつセンチ単位のオーダーで安定的に対象物の空間的な位置を検出することが可能な超音波方式を採用する。

本研究では、超音波による空間位置センサとして超音波三次元位置測定システム (Zone Positioning System, 古河機械金属株式会社製) を使用する。実験環境は約4m×4m×2.7mの空間であり、その周辺は高さ1.9mの遮光パネルで囲まれている。天井には受信機 (以下、リーダ) がほぼ格子状に25箇所設置されている。発信源である超音波送信機 (以下、タグ) は50mm×50mm×15mmで30g (電池装着時) と小型かつ軽量なものであり、発信源を上面と側面の2箇所に搭載することで、タグの向きに対するロバスト性が改善されている。また、リーダ、タグ共に100°の円錐状で広がる指向性特性をもっている。

#### 3. 3 検出対象

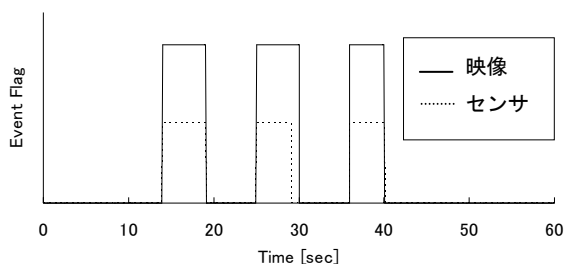
検出するイベントとして除細動を対象とする。除細動は心停止の状態に対して、電気的なショックを与え生還に導くきわめて重要な処置であり、このイベントを検出することは事後検証の注目箇所として利用されることが期待される。

#### 3. 4 検出方法

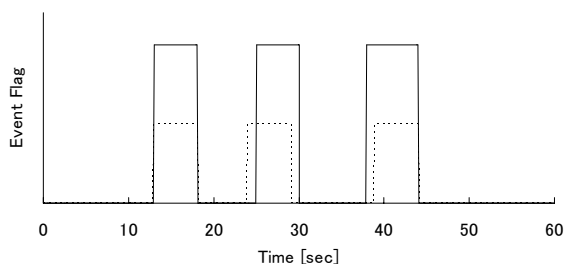
本研究では、除細動器のパッドの三次元位置と動作時間がある閾値以上であれば、除細動器によって除細動が行われていると仮定し、2つのパッドに装着したタグから計測される位置情報に基づいてイベントを検出する。具体的には、ベッドに横たわる患者の胸部空間領域と除細動器の停留時間を予め設定し、除細動器の空間的軌道が設定領域と設定停留時間の条件を満たせばイベントの開始時刻を、条件からはずれれば終了時刻を記録する。除細動検出フローを図5に示す。



図6 除細動を実施している様子



(a)動作パターン1



(b)動作パターン2

図7 実験結果

### 3. 5 実験

本手法の有効性を示すため、除細動の処置動作を模倣し、センサ情報による除細動検出の可能性を検証するために実験を行った。図6に除細動の様子を示す。除細動器の場所や処置者の姿勢に関係なくイベントが検出できることを確認するため、実際の治療現場で想定される動作、つまり(a)処置者は患者の右側から直立した姿勢で処置する場合と、(b)処置者は患者の左側から覆い被さった姿勢で処置する場合の2通りの動作パターンについて3回の除細動を行った。また、映像を目視で判断した除細動実施区間と本手法で得た検出結果が一致しているかを評価するために、天井に設置したカメラから処置者の様子を撮影した。

実験結果を図7に示す。横軸は時間、縦方向はイベントフラグを示す。つまり除細動が行われていない区間をOFF、除細動が行われている区間をONとする。いずれの動作パターンも3回の除細動を検出しており、本手法によりイベントを有効に検出していることが確認できる。また、三次元位置情報を利用したイベント検出結果と実際にイベントが行われた時間との間に1秒程度のズレが認められるが、映像を目視によって除細動実施区間を判断する際の誤差によるものと考えられる。

### 4. まとめ

本研究では映像によるイベント検出として、心臓マッサージのイベントをCPR映像から自動的に検出する手法に関する検討を行った。また、映像情報を用いず、超音波三次元タグを用いて除細動のイベントを検出する手法に関しても同様に検討を行った。その結果、二つの手

法ともほぼ正しくイベント検出できることがわかった。しかしながら、映像によるイベント検出においては、対象領域設定の問題が残されている。今後は、現在手動で行っている対象領域の設定を、画像認識技術などの活用により自動化することを検討する。また、救急救命処置で行われる各種の処置行為と三次元位置の関係の検討を進め、リアルタイムに医療行為を大別することを検討する予定である。それぞれの利点を活かし、二つの手法を同時に用いることで精度良くイベント検出を行う手法等についても検討を行いたい。

### 謝 辞

本研究は文部科学省知的クラスター創成事業の一環としてなされたものである。本研究を遂行するにあたりシステム評価の場を提供いただいたACLS岐阜の方々、多くの有益な助言をいただいた岐阜大学工学部速水研究室、岐阜大学大学院医学系研究科救急・災害医学分野及び三洋電機株式会社の方々に深く感謝する。

### 文 献

- [1] 浅井博次, 棚橋英樹, 速水悟, “診療記録システムの開発”, 岐阜県生産情報技術研究所研究報告, No.7,pp.5-7,2006.
- [2] AHA心肺蘇生と救急心治療のためのガイドライン 2005, ISBN4-521-67701-0
- [3] H. Asai, H. Tanahashi, et al., "The Cardiac Massage Detection in The Emergency Medical Care Video", Proc. of the Sixth IASTED International Conference on Visualization, Imaging, And Image Processing, pp.597-602, 2006