

(註)本稿は(社)日本映画テレビ技術協会機関誌「映画テレビ技術」No.649(2006年9月)に掲載された同タイトル記事の最終ドラフトですので、取り扱いにご注意下さい。  
著者: 柴田賀昭

## エッセンスマーク TM

### はじめに

昨年末、SMPTE(全米映画テレビジョン技術者協会)の機関誌である SMPTE Motion Imaging Journal の 2005 年 12 月号に"EssenceMark - SMPTE Standard-based Textual Video Marker -"なるタイトルで「エッセンスマーク」の紹介記事を掲載しましたところ非常に好評を頂き、多方面から様々な反応が寄せられました。

そのひとつとして今回、本誌へ「エッセンスマーク」の日本語での解説記事を執筆する機会を与えて頂きました。ただオリジナルの記事は元々技術論文ということで、「エッセンスマーク」に関連した SMPTE 標準技術の詳細を始め、「エッセンスマーク」対応を実装するのに必要な技術情報をこと細かく説明したりしましたので、これをそのまま翻訳してしまうとかなり堅苦しい話になってしまいそうです。

そこで本稿ではそのような技術詳細はあえて避け、厳密性よりはむしろ判りやすく説明することを心掛けて、「エッセンスマーク」のエッセンス(!?) を紹介してみたいと思います。

### 「エッセンスマーク」とは

「エッセンスマーク」を説明する上で筆者がいつも用いる例えは「ポストイット」(米 3M 社の登録商標)です。ご存知の通り、「ポストイット」とは“簡単に引き剥がしができる糊のついた小さな紙切れ”でして、例えば本のしおり代わりに使ったり、分厚い書類の(後で見直したい)興味ある部分への仮マーキングに用いたりすることは多くの方々に馴染みのあるところかと思えます。

さらに「ポストイット」には様々な大きさやカラーバリエーションが準備されていますから、目的や興味の違いに応じてこれらをうまく使い分けることが可能ですし、あるいは仮に手元にある「ポストイット」が一種類であれば、目的や興味を文字にして「ポストイット」自体にペンで直接書き込むことでそれを使い分けるといった方法もあります。

「エッセンスマーク」とは、まさにこの「文字入りポストイット」を映像音声データ(以下、AV 素材あるいは単に素材と呼びます)に適用したものです。

図 1 に、「エッセンスマーク」が付与された AV 素材の一例を模式的に示します。

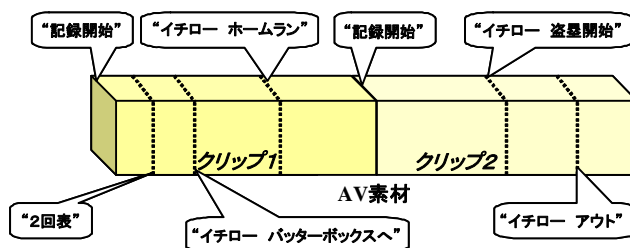


図 1: 「エッセンスマーク」が付与された AV 素材

図 1 では、あの有名なシアトル・マリナーズのイチロー選手を撮影した AV 素材に対して、後段の編集で使いそうな部分に「エッセンスマーク」を付与した事例を示しています。例えば最初のビデオクリップ(クリップ 1)は、2 回表においてイチロー選手がホームランを打ったシーンを撮影したものであり、そのクリップの 2 回表の開始位置、イチロー選手がバッターボックスに入った位置、そしてイチロー選手がホームランを打った位置に対して、それぞれ“2 回表”、“イチロー バッターボックスへ”、“イチロー ホームラン”なる値をもった「エッセンスマーク」を打ち込んでいます。

これらの「エッセンスマーク」は、撮影後の例えば AV 素材のレビューあるいはロギングなどといった工程で付与される場合が多いと思われませんが、適切な機器、例えば撮影中のカメラに連動して「エッセンスマーク」を打ち込むことができる携帯機器などを用いることで、実際の撮影に併行して「エッセンスマーク」をその場でリアルタイムに打ち込んでいくといった方法なども想定されます。

またここで示した「エッセンスマーク」は、主として人間が自らの判断でもってマニュアル操作で挿入することを想定していますが、その一方で、例えば 図 1 中にて各ビデオクリップの開始位置を示した“記録開始”のように、機器が何らかのデータ処理の結果として「エッセンスマーク」を自動的に生成、挿入していくといった場合も想定されます。

しかしながら「エッセンスマーク」の仕様自体としては、それらの生成方法については何ら規定しておらず、後述するように単に AV 素材に対する電子的な「文字入りポストイット」としての枠組みを規定しているに過ぎません。

ところで、「エッセンスマーク」の名前のいわれですが、ここで言う「エッセンス」とは“本質的なもの”ではなく“AV 素材”そのものを表す SMPTE 用語です。デジタルの世界においては AV 素材もメタデータも共に 0、1 の並びに過ぎませんが、やはり両者には本質的な違いが存在する訳ですから区別が必要。そこで SMPTE では（従属データとしてのメタデータに対して本質的なデータである）AV 素材のことを「エッセンス」と呼んでいるものと思われま

### 「エッセンスマーク」の特長

さて、これまでの「エッセンスマーク」の説明から、世の中には既に似たようなメタデータが多数存在しているのでは？とお感じになった読者も多いかと思われま

す。例えば素材収録後のレビューにおいて、着目すべきフレームの内容や理由をそのタイムコードと共にノートに記録していくなどといった作業は、まさに「エッセンスマーク」と同様のメタデータを手動で生成しているとみなすことができます。

また業務用、民生問わず多くの映像機器において、映像データの所望の位置にマーキングを施し、あとでその部分の高速読み出しを可能とするといった機能がサポートされています(例えば弊社の業務用 VTR の場合、「グッドショットマーク」といった名称で当該機能をサポートして参りました)。

そもそも膨大なデータを扱わなければならない状況において後で振り返るために所望の位置にマーキングを施しておくといった行為は、例えばウェブブラウザにおけるブックマーク機能を例に挙げるまでもなく、ごく基本的な要求です。そしてこのような要求に応えるべく用いるメタデータを、ある着目“点”にマーキングするものということで、POI (Point Of Interest)メタデータと呼んだりします。

そして先述したように、様々な映像機器あるいはアプリケーションが何らかの形で POI メタデータをサポートしている訳ですが、ここで問題になるのはそれらの間の相互運用性です。つまり現状では、ある特定の映像機器、アプリケーションあるいはそれらの関連製品群の間での POI メタデータのやり取りは可能であっても、不特定多数の映像機器やアプリケーションの間でそれを実現することは非常に困難であり、その結果、例えば折角長時間を掛けて収録素材をレビューし重要箇所をマーキングしても、それを後段の編集工程に自動的にもって行く術がなく、結局はレビュー時のノートの手書き記録に頼らざるを得ないといった状況に陥ってしまいます。

「エッセンスマーク」はこのような問題を解決するために開発されました。すなわち様々な映像機器やアプリケーションの間での POI メタデータのやり取りを実現すべく、「エッセンスマーク」は業界標準である SMPTE メタデータ仕様をその土台とし、これに詳細な運用ルールを追加仕様として取り決めることで生み出されました。これより、仮に対象となる映像機器あるいはアプリケーションが「エッセンスマーク」のことを知らずとも、後述する業界標準としての SMPTE メタデータの取り扱いをサポートしていれば、それは自ら「エッセンスマーク」もサポートしていることとなります。次の節ではこの SMPTE メタデータ仕様について、その触りの部分を紹介したいと思います。

### SMPTE メタデータと「エッセンスマーク」

SMPTE メタデータ仕様の最も基本的な考え方は、全てのメタデータを、その項目を表すラベル(例えば“筆署名”)と実際の値(同、“柴田”)との組み合わせとして表現しようというものです。具体的には、図 2 に示すように、全ての SMPTE メタデータは KLV パッケージと呼ばれる Key-Length-Value なる構造をもったバイト列として表現されています。

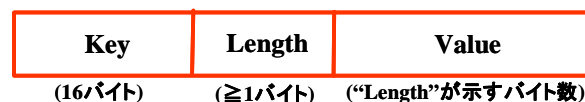


図 2: SMPTE KLV パッケージ

図 2 において Key は SMPTE によって決められたメタデータ項目を表すラベルであり、16 バイトデータとして表現されます。別の言い方をすれば、SMPTE は様々なメタデータ項目を、それに割り当てた 16 バイトデータと共にメタデータ辞書と呼ばれる文書によって一元管理しており、その登録総数は既に 2000 項目近くに及んでいます。

また Value は、Key で指定されたメタデータ項目に対応した実際の値であり、それに先行する Length によって、Value のデータ長が前もって示されています。

そして「エッセンスマーク」の実体と言え、それは“Term Value”と呼ばれるメタデータ項目を表すラベルを Key とし、付与したい所望の文字列を Value に格納した KLV パッケージそのものに過ぎません。ただし、より広範囲での相互運用性を達成すべく、「エッセンスマーク」としては、例えば Value は最大 32 バイトに限定するとか、日本語などマルチバイト文字の格納方法を厳密に規定するなどの運用ルールを追加仕様として定めています。

このように「エッセンスマーク」は、言わば SMPTE メタデータのひとつの応用例として提供されていることから、SMPTE メタデータの取り扱い方法として規定された全ての技

術仕様は「エッセンスマーク」にもそのまま適用可能です。例えば SDI(Serial Digital Interface)を介して AV 素材をベースバンド伝送する場合、KLV パケットは VBI(Vertical Blanking Interval)領域に補助データとして格納するといったルールが定められていますので、「エッセンスマーク」もまた、それが参照するフレームの VBI 領域に格納して伝送されることとなります。

また最近ホットな話題となっている MXF (Material Exchange Format)に基づいたファイルでの素材のやり取りの場合も同様です。すなわち MXF ファイルのボディ部において、SMPTE メタデータを AV 素材に重畳して伝送、記録する方法が規定されていますので、これらをそのまま「エッセンスマーク」にも適用すれば良いということになります。

### 「エッセンスマーク」の拡張機能

これまで述べたように、「エッセンスマーク」として規定したのは、AV 素材の所定の位置に任意の文字列を付与するための枠組みに過ぎません。しかしこれだけですと、様々な映像機器やアプリケーションの間で情報をやり取りするには心許ない部分があります。

そこで更に相互運用性を高めるべく、「エッセンスマーク」仕様の一部として、幾つかの「エッセンスマーク」値を、その解釈方法も含めて“予約”することとしました。ここで予約とは、「エッセンスマーク」値の表記及び解釈方法を予め定義しておくことです。

図 3 は本稿の執筆時点で予約している「エッセンスマーク」値を、その解釈方法と併せて一覧表にまとめたものです。

予約語	解釈方法
_RecStart	AV 素材記録の開始点
_RecEnd	AV 素材記録の終了点
_ShotMark1	着目点 1
_ShotMark2	着目点 2
_Cut	カット点
_Flash	フラッシュ光の検出点
_FilterChange	フィルタの変更点
_ShutterSpeedChange	シャッタ速度の変更点
_GainChange	ゲインの変更点
_WhiteBalanceChange	ホワイトバランスの変更点
_OverBrightness	ビデオ信号の出力レベルが 100%を超えた点
_OverAudioLimiter	オーディオ信号の出力レベルが所定の閾値を超えた点
_In-XXX	イン点(XXX は任意の文字列)
_Out-XXX	アウト点(同上)
_KeyFrame	キーフレーム

図 3: 「エッセンスマーク」予約語一覧

これより、例えば撮影中にカメラマンが着目した点に“\_ShotMark1”なる「エッセンス

マーク」を付与しておけば、それは後段の、例えば収録素材の着目点の一覧表示といったアプリケーションに引き継がれることとなります(後で具体例を示します)。

さらに「エッセンスマーク」予約語を別の側面から見れば、それは様々な映像機器及びアプリケーションに対して「エッセンスマーク」値の解釈方法、つまりその“意味”を共有させていくことへとつながります。

例えばいま編集済み素材のカット点検出が可能な映像再生機を想定し、そのような映像再生機に対して再生時に検出したカット点位置に“\_Cut”なる「エッセンスマーク」を自動挿入する機能を追加したとします。すると、この再生映像信号を受け取って記録する映像記録機が“\_Cut”「エッセンスマーク」を解釈可能であれば、その記録機に例えば“\_Cut”が付与された位置を基準に入力映像信号を分割ファイルとして保存するといった動作をさせることができます。これはまさに、それらの機器が“\_Cut”なる文字列を共通言語として“会話”し、それが意味するところに応じた処理をおこなっているとみなすことができます。

あるいは特に映像記録機においては、入力映像信号を記録するにあたり、その記録機器に入力映像信号を分割ファイルとして保存させるための制御コマンドとして“\_Cut”「エッセンスマーク」を利用している、といった見方をすることもできます。

#### 「エッセンスマーク」応用の実際例

それでは「エッセンスマーク」が映像制作の現場で具体的にどう用いられているか、ここでは幾つかの例を弊社製品への実装例も交えながら紹介したいと思います。

まず先にも少し触れましたが、「エッセンスマーク」の生成方法は、(a)人間が自らの判断に基づいて生成するもの、(b)映像機器あるいはアプリケーションが自動生成するもの、のいずれかに大別されます。これらを模式的に図 4 に示します。

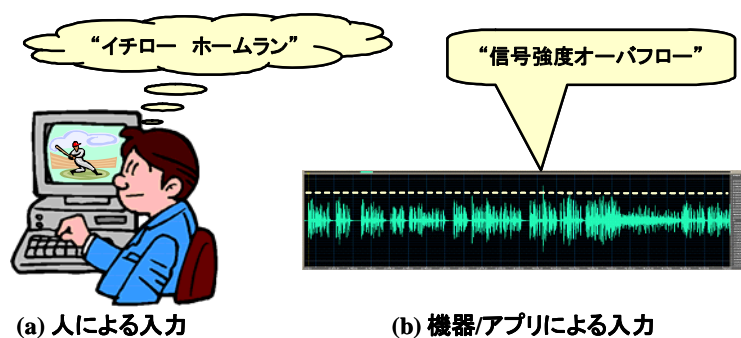


図 4: 「エッセンスマーク」の生成方法

図 4(a)では、野球の収録素材(あるいはライブ中継時でも構いません)のレビューにおいて、イチロー選手がホームランを打ったシーンを検出し、そこに“イチロー ホームラン”なる「エッセンスマーク」を付与した様子を示しています。図 1 に示した「エッセンスマーク」付き AV 素材は、まさにこのような作業を通じて生成されるものです。

なお、このような「エッセンスマーク」入力においては、もちろん入力すべきと判断したその場で所望の語句、例えば“イチロー ホームラン”を入力、付与しても構いませんが、例えば“イチロー バッターボックスへ”、“イチロー ホームラン”、あるいは“イチ

ローアウト”など入力が見定される語句をそれらを割り当てたボタンと共に予め登録しておき、レビューの際には所定のボタンを押すことで、それに割り当てられた語句を「エッセンスマーク」値として直ちに付与するといった方法が有効になると思われます。

ところで、弊社が提供する業務用カムコーダの多くは、以前から「グッドショットマーク」なる名称で映像データの所望の位置にマーキングを施すといった機能を提供して参りましたが、現在ではこれと互換性を持たせたかたちで、着目点を表す「エッセンスマーク」である“\_ShotMark1”あるいは“\_ShotMark2”（図 3 参照）を同時に生成、付与するといった機能も搭載しております。

さて、図 4(a)に対し同(b)は、例えば入力オーディオ信号をモニタリングし、その信号強度が所定の閾値を越えたところを自動的に検出して「エッセンスマーク」でマーキングするといった様子を示しています。このような処理によって、例えば爆発のシーンや拳銃の発砲シーンなどを自動的に検出しマーキングを施すことが可能となります。

このような方法は単にオーディオ信号のみならず映像信号でも有効です。例えば先に紹介したカット点検出による“\_Cut”「エッセンスマーク」の自動挿入機能は、まさに映像信号の解析結果を「エッセンスマーク」で表現したものとみなすことができます。

現在、弊社でもこのような「エッセンスマーク」の自動生成機能のサポートに取り組み始めております。具体的には弊社の最近の業務用カムコーダ、例えば次世代光ディスクを用いた映像機器である XDCAM™ シリーズのカムコーダの一部の機種(PDW-510/530)では、先述した着目点を表す「エッセンスマーク」に加え、“\_Flash”、“\_ShutterSpeedChange”、“\_WhiteBalanceChange”、“\_OverAudioLimiter”といった「エッセンスマーク」の自動生成、付与機能を搭載しており、その結果、後述する「『エッセンスマーク』に基づくサムネイル表示機能」と併せて、収録素材の特徴部分を簡単に頭出しできるようなアプリケーションを提供しております。

次に「エッセンスマーク」の利用法ですが、先述の“\_Cut”「エッセンスマーク」を用いた例のように、予め決まった値を持つ「エッセンスマーク」をあたかも機器の制御コマンドとして用いるなどといった高度な使い方もありますが、最も基本的な使い方としては、やはり「エッセンスマーク」が付与された位置の映像を簡単に見るための手段を提供するといったことでしょうか。例えば図 5 は「エッセンスマーク」に関する XDCAM ユーザーインターフェースの一例を示したのですが、ここでは光ディスクに記録した全ての素材において、“ShotMark1”「エッセンスマーク」（図 3 の“\_ShotMark1”に相当）が付与されたフレームをサムネイル化して一覧表示しています。このような機能を通じて、撮影時あるいは収録素材の事前レビュー時に“着目点 1”としてマーキングした部分をまとめて確認することができますし、また所望のサムネイル画の選択を通じて、ある着目点への高速頭出しや自動再生なども簡単におこなうことができます。



図 5: 「エッセンスマーク」に基づくサムネイルリスト

これを更に発展させると、いわゆる“着目点1”をテーマにしたビデオサマリ(要約)を自動的に合成するといったことも可能となります。図 6 はその様子を概念的に示したものです。いま、例えば“AV 素材 1”、“AV 素材 2”なる二つの収録素材があり、それらに対して所定の“着目点1”に基づいて“ShotMark1”「エッセンスマーク」が計三箇所が付与されていたとします。この時、“着目点1”をテーマとしたビデオサマリとは、“ShotMark1”「エッセンスマーク」が付与されたフレームを含むかたちで例えば5秒の区間を切り出し、これらを収録時間順に並べて計15秒のビデオクリップとしたものです。

このようなビデオサマリは、テーマとなる「エッセンスマーク」値と指定区間の切り出し条件さえ与えてやれば自動的に合成することができます。別の例としては、例えばいま有名人の記者会見を“\_Flash”「エッセンスマーク」(図 3 参照)を自動的に挿入しながら収録した素材があったとします。この場合、テーマとして“ShotMark1”の代わりに“\_Flash”「エッセンスマーク」を選択すれば、その有名人に対してストロボが焚かれた瞬間、つまり多くのカメラマンが着目した瞬間をまとめたビデオサマリを自動的に合成することができる訳です。

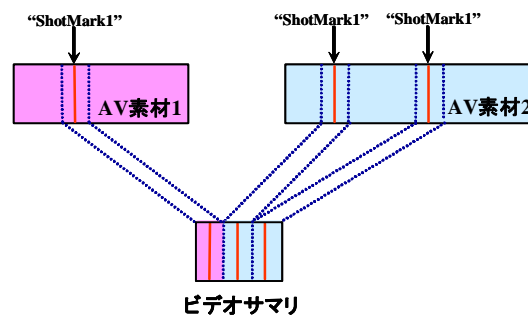


図 6: 「エッセンスマーク」に基づくビデオサマリ

ここで紹介した、着目点のサムネイル画の一覧表示やビデオサマリなどといったアプリケーションは決して目新しいものではありませんが、その基本となる部分に「エッセンスマーク」を用いることで様々な横展開が可能となります。すなわち先に紹介したサムネイル画の一覧表示は単に前出の XDCAM カムコーダで撮影した素材のみならず、従来の「グ

ッドショットマーク」対応のカムコーダで撮影したテープ素材を XDCAM ディスクへダビングした素材でも実現することができます。

また、弊社が提供する一部の VTR には再生時にカット点を検出して “\_Cut” 「エッセンスマーク」を自動挿入するといった機能を搭載したものもありますが、これらと併せて用いることで、例えば過去の完パケの再利用を目的にショット分割をおこなうための手掛かり(完パケ構成ショットのサムネイル画による一覧表示)を与えたりすることもできます。

### まとめに代えて

ある調査によれば、映像制作において最も時間を要する工程は収録した素材から所望の箇所を選び出す素材の検索及び切り出し工程であり、場合によっては全作業時間中の半分近くをそれに費やしてしまうといった事例もあるそうです。

もちろん本当に良い部分を選択するには制作者の創造的な判断に委ねるしかない訳ですが、他方で素材のデータ量がますます膨大になるにつれ、所望の位置へ何度も簡単にアクセスしたいとか、とりあえずの選択候補を自動的に示して欲しいといった要求は、限りある制作時間をより創造的な作業へ割り当てる上で非常に基本的なものだと思われま

す。「エッセンスマーク」には、機器あるいはアプリケーションを問わずこれを解決できる可能性があります。これより読者におかれましては、まずは「エッセンスマーク」に慣れ親しんで頂いて日頃の映像制作業務の効率化を狙い、また新たな要求を元に「エッセンスマーク」の活用方法を我々にご提案頂ければ幸いに存じます。また冒頭で申しましたように「エッセンスマーク」の詳細な技術仕様自体は既に SMPTE 機関誌の方で公開していますので、もし技術詳細に興味を持って頂けましたら是非こちらの論文も併せてご覧頂ければと思います。

そしてまさに、書籍やホームページの整理には「ブックマーク」があり書類整理には「ポストイット」があるように、AV 素材の整理のための基本ツールとしての「エッセンスマーク」が業界で広く受け入れられ、読者の創造的作業にいささかでも貢献することができれば、筆者としては望外の喜びです。

以上