

Development and Expansion of Video Pen System

Fuma Kishi †, Yuki Miura † and Takayuki Shinoda †

Summary: With the aim of streamlining on-site operations and realizing new program expressions, we have in-house developed a video pen system that allows anyone to easily use new technologies such as AI. The system allows for intuitive operation of functions such as "real-time player tracking CG," "automated player name CG," and "on-field distance angle estimation." Furthermore, by enclosing information on PC screens, etc., we have achieved "automated score data entry" and "automated relay race course guides," thereby improving the efficiency of many program production scenes. Furthermore, we have planned a "segment where AI calculates the excitement of the game in conjunction with the game timer," providing a new dimension to sports highlight scenes.

Keywords: AI, image recognition, television production, graphics

1. Introduction

In recent years, TV stations have been actively using AI to reform work styles and improve broadcast content.

Although many innovations have been made in the output of information and CG when using computers, many of the operating interfaces have been clunky and difficult to use.

Therefore, we focused on a system called the Video Pen, which has the advantage of being intuitive to operate, and redeveloped it into a system that anyone can easily use, including functions such as image recognition. In this paper, we report on the specific functions and operational results.

2. Key points of the initiative

Nippon Television has been developing a variety of technologies, including AI, and has now integrated these various functions to develop a "multifunctional device that anyone can use easily." This system processes information within the video by touching or surrounding it with the freehand on input video such as SDI, and outputs the results as data or CG. Figure 1

A simple overview of this system is shown below. Note that an operation tablet is not necessarily required, and operation with a normal mouse is also possible. As such, the system is very simple and easy to set up. Therefore, this system does not necessarily require the preparation or presence of a technical person with specialized knowledge.

Its high

operability and ease of setup have led to a wide range of content, leading to new expressions and more efficient program production.

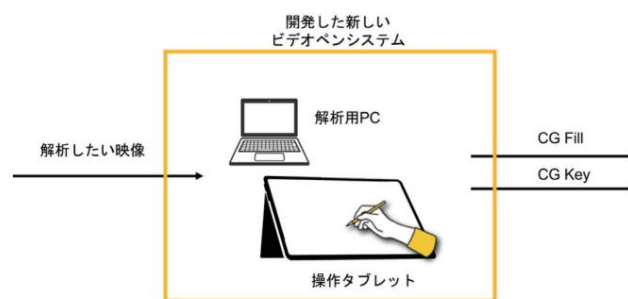


Figure 1 System diagram

This system is also characterized by the flexibility to add new functions, which is possible only with in-house development, and we can expect to continue adding new functions based on various new technologies in the future.

3. Functions of the newly developed video pen

This chapter introduces the use cases of the new video pen we have developed. While doing so, we will explain the functions of the system.

3.1 Touch function

First, we will introduce the basic function, the touch function. 3.1.1

Real-time positioning In live sports broadcasts,

many players often appear on screen and play simultaneously. For example, in soccer, 22 players play on the court at the same time. For this reason, during editing, in order to draw attention to a particular player, player-following CG is sometimes displayed by superimposing CG that follows the player in the footage. However, like mosaic processing, this is a time-consuming process, making it difficult to use in live broadcasts.

Therefore, this system combines image recognition and touch functionality.

It automatically follows the object you touch in real time.

[Copyrights to Machine-Translated Content]

The copyright of the original papers published on this website belongs to the Institute of Image Information and Television Engineers. Unauthorized use of original papers or translated papers is prohibited. Please be sure to cite the original publication when referencing. For details, please refer to the copyright regulations of the Institute of Image Information and Television Engineers.

ビデオペンシステムの開発と機能拡充

Development and Expansion of Video Pen System

岸 楓馬[†], 三浦 祐樹[†], 正会員 篠田 貴之[†]

Fuma Kishi[†], Yuki Miura[†] and Takayuki Shinoda[†]

あらまし 現場オペレーションの効率化や、新しい番組表現の実現を目的として、AI等の新技術を誰でも簡単に扱えるビデオペンシステムを社内開発した。「リアルタイム選手追従CG」、「選手名CGの自動化」、「フィールド上の距離角度推定」などの機能を直感的な操作で実施可能なシステムである。またPC画面等の情報を囲むことで「得点データ入力 of 自動化」、「駅伝コースガイドの自動化」を実現させ、多くの番組制作シーンでの業務効率化を行った。さらに「試合のタイマと連動しAIが試合の盛り上がりを算出するコーナー」を企画し、スポーツのハイライトシーンに新たな軸を提供した。

キーワード: AI, 人工知能, 画像認識, 番組制作, グラフィック

1. ま え が き

近年、テレビ局では働き方改革や放送内容の高度化のためにAIを積極的に用いた取り組みが盛んである。AIを用いる際、これまで情報やCGの出力には多くの工夫がなされてきたものの、その操作インタフェースに関しては無骨で扱いにくいものが多く存在していた。

そこでわれわれは直感的な操作感が強みであるビデオペンというシステムに着目し、画像認識などの機能を誰でも簡単に扱うことができるシステムへと再開発した。本稿では具体的な機能や運用実績について報告する。

2. 取り組みのポイント

これまで日本テレビではAIをはじめとするさまざまな技術開発を行ってきたが、これらの各種機能を統合し、「多機能でありながら誰でも簡単に扱えるデバイス」を開発した。本システムはSDI等の入力映像に対して、フリーハンドでタッチもしくは囲むことで映像内の情報を処理し、その結果をデータやCGとして出力するシステムである。図1に本システムの簡単な概要を示す。なお必ずしも操作タブレットが必要なのではなく、通常のマウスでの操作も可能である。このようにシステムが非常にシンプルであり、セットアップも簡単である。そのため本システムは専門知識を持った技術担当者の準備や立ち合いを必ずしも必要としない。

その操作性の高さやセットアップの容易さからさまざまなコンテンツへ広がりを見せ、新しい表現や番組制作の効

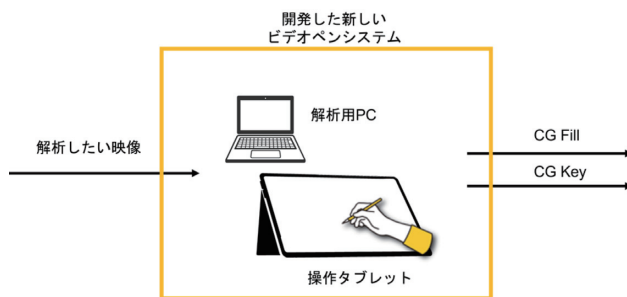


図1 システム図

率化に貢献した。また本システムは社内開発ならではの機能追加の柔軟さが特徴であり、今後もさまざまな新技術の機能追加が期待できる。

3. 開発した新しいビデオペンの機能

本章では開発した新しいビデオペンの活用ケースを紹介しながら、システムが有する機能について説明する。

3.1 タッチ機能

初めに基本的な機能であるタッチ機能を紹介する。

3.1.1 リアルタイムポジショニング

スポーツ中継では、多くの選手が一度に画面に映り同時にプレーすることが多い。例えばサッカーではコート上で22人が同時にプレーする。そのため編集時には、特定の選手に注目させるため、映像中の選手に追従するCGを重畳させる選手追従CGの表示を行う場合がある。しかしモザイク処理と同様に手間がかかる作業であり、生放送で活用するのは困難である。

そこで本システムでは画像認識とタッチ機能を組み合わせ、リアルタイムでタッチした対象を自動で追従し続ける



Figure 2 Operation scene



Figure 3. Defense line display

It was developed to be able to display CG. When you touch a player, the CG that follows that player is displayed in real time. The operation is shown in Figure 2. Furthermore, when you touch multiple players, it is also possible to display lines with the shortest distance between the players, starting from the top of the screen. For example, in soccer, as shown in Figure 3, the defensive line can be displayed.

When touching multiple players, you can register the uniform colors of each team by touching each player on both the opposing and friendly teams on the day. This prevents lines from connecting even if you accidentally select different teams. In addition, even if players switch positions while the defensive line is displayed, the system constantly recognizes the positions of the players and automatically adjusts the relative positions of the lines to prevent lines from crossing.

Furthermore, as another function, it is possible to select the caption you want to display for the target and display it while tracking. For example, in a baseball broadcast, after selecting a position tag such as first, second, or third, you can touch the player and have the position CG displayed above his head in real time.

3.1.2 Real-time player name CG

This system not only recognizes objects, but also recognizes faces.

It also has a built-in recognition function, which makes it possible to clearly display information about athletes during live sports broadcasts. For example, more than 30,000 people participate in the Tokyo Marathon, with many athletes sharing the same last name and many from overseas. This makes it extremely difficult to instantly determine who is appearing on the screen during a live broadcast and display their names. Additionally, because athletes run in groups, viewers have a hard time identifying at a glance which athletes are attracting attention. To

address this issue, this system has two types of athlete name display functions.

The first mode allows you to specify a player's name on the UI and displays a caption for that player. This is used when you want to focus on just one specific player.

The second mode is to recognize the faces of all players who have been registered in advance and automatically display the player's name CG. This is useful when you want to show many notable players.

However, road racing

is one of the sports where facial recognition is difficult because the expressions of athletes change due to fatigue and the color and brightness of sunlight changes as the race progresses.

This system has succeeded in improving the accuracy of facial recognition during a race by continually updating the faces of athletes who have been recognized even once during the race. In actual competitions where it has been used, accurate display has been achieved without a single mistake. Furthermore,

when multiple CG player names are automatically displayed, there is a concern that the CG player names may unintentionally overlap. Therefore, the system is designed to recognize the top and bottom and size of the athletes' faces and automatically adjust the position so that the CG player names do not overlap. For this reason, they are not necessarily displayed directly above the athletes, but if athletes are crowded together, the display position is automatically moved so that the CG does not overlap. Taking this into consideration, if the display overlaps, the system estimates the relative positions of the riders from the size of their faces and displays only the rider in front.

This type of ingenuity makes it possible to prevent mistakes in advance, even under the extremely harsh conditions of a road race. It also means that the system can be operated without the need for a dedicated operator, leading to labor savings. Furthermore, because rider names are displayed automatically, it also helps to reduce the number of mistakes made by announcers when reading them.

3.1.3 Range and angle estimation

In rugby, kicking is a big highlight.

In rugby, CG images of the distance and angle between the ball and the goal are often displayed to indicate the difficulty of the kick. Until now, distance measurements were carried out on-site in advance, and a model of the entire rugby field was created to calculate the distance and angle. Furthermore, since the relative positions of the field and camera differ at each venue, measurements had to be taken in advance for each match, making preparations difficult. This system uses the white lines drawn on the field as a clue and can estimate the distance between the ball and the goal from video alone. In rugby, there are several white lines with specified distances, such as the 5m line and the 22m line. These white lines

[Copyrights to Machine-Translated Content]

The copyright of the original papers published on this website belongs to the Institute of Image Information and Television Engineers. Unauthorized use of original papers or translated papers is prohibited. Please be sure to cite the original publication when referencing. For details, please refer to the copyright regulations of the Institute of Image Information and Television Engineers.



図2 操作風景

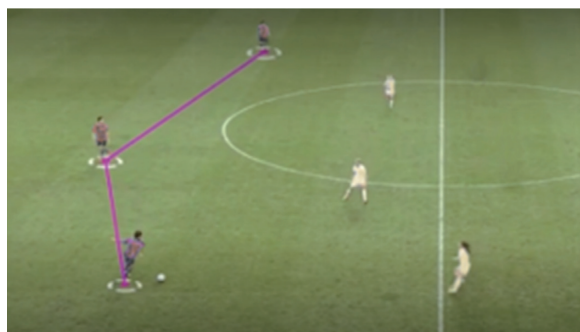


図3 ディフェンスライン表示

CGを表示できるよう開発した。選手をタッチすると、その選手に対し選手追従CGをリアルタイムで表示する。操作の様子を図2に示す。さらに、複数選手をタッチすると、画面上部から順に選手間の最短距離でラインを表示することも可能である。例えば図3のようにサッカーにおいてディフェンスラインを表示できる。

複数選手をタッチする際、当日に敵・味方チームの選手を一人ずつタッチすることで、各チームのユニフォームの色を登録できる。これにより誤って異なるチーム同士を選択した場合でもラインを繋がないように工夫している。また、ディフェンスライン表示途中に選手の上下が入れ替わってもラインがクロスしないように選手の位置を常に認識し続けラインの位置関係を自動調整するようにしている。

さらに、別の機能として、対象に対して表示したいテロップを選択し、トラッキングしながら表示することも可能である。例えば野球中継においてファースト、セカンド、サードといったポジションのタグを選択した後、選手をタッチすることで頭上にポジションCGをリアルタイムで表示できる。

3.1.2 リアルタイム選手名CG

本システムでは先ほど述べた物体認識だけでなく、顔認

識機能も搭載しており、スポーツ中継において選手の情報をわかりやすく表示できる。例えば東京マラソンでは3万人以上が参加し、同姓の選手や外国人の選手も多い。そのため生放送中に画面に映っている選手が誰かを瞬時に判断し、選手名を表示することは非常に難しい。また選手がかたまりとなって走るため、視聴者にとっても注目選手が誰であるか一目でわかりづらいという問題も発生していた。

そこで本システムでは2種の選手名表示機能を有する。

一つ目はUI上で選手名を1人指定し、注目する選手1人のテロップを表示するモードである。こちらは特定の選手1人だけに着目したい際に使用されている。

二つ目は事前に登録しておいた選手全員に対し、画面に映る選手の顔を認識し、自動で選手名CGを表示するモードである。こちらはたくさんの注目選手を見せたい場合に使用されている。

しかし、疲労によって選手の表情が変わったり、競技の時間経過によって太陽光の色や明るさが変わることから、ロードレースは顔認識が難しい競技の一つである。そこで本システムでは一度でも顔認識した選手の顔はレース中も逐次更新し続けることでレース中の顔認識の精度を向上させることに成功した。実際運用した競技において一度も間違えることなく正確な表示が実現できている。

また、複数の選手名CGを自動表示する場合、意図せず選手名CG同士が被ることが懸念される。そこで選手の顔の上下や大きさを認識し選手名CGが被らないように自動で位置を調整するよう工夫した。そのため必ずしも選手の真上に表示するわけではなく、選手が密集している場合にはCGが被らないように自動で表示位置を移動させる。それをふまえた上で表示が被ってしまう場合は、顔の大きさから選手の前後の位置関係を推定し、前方の選手のみを表示する。

このような工夫により、ロードレースという非常に厳しい条件下であっても、事前にミスを防ぐことができる。またこれにより専用のオペレータを配置することなく運用できるため、省人化につながる。また選手名が自動で表示されるためアナウンサの読み上げミスを削減するための手助けにもなった。

3.1.3 距離角度推定

ラグビーにおいて、キックは大きな見せ場である。そこで、キックの難易度を示すためにゴールとボールとの距離や角度のCGを表示する場合が多い。これまでは、あらかじめ現地で距離計測などを行い、ラグビーフィールド全体のモデルを作成し距離や角度を算出していた。また、会場ごとにフィールドとカメラの位置関係が違うので、毎試合事前計測が必要となり準備が大変であった。本システムではフィールドに引かれている白線を手掛かりに、映像のみからボールとゴールまでの距離を推定することが可能である。ラグビーには、5mラインや22mラインなど、いくつか距離が規定されている白線が存在する。そういった白線

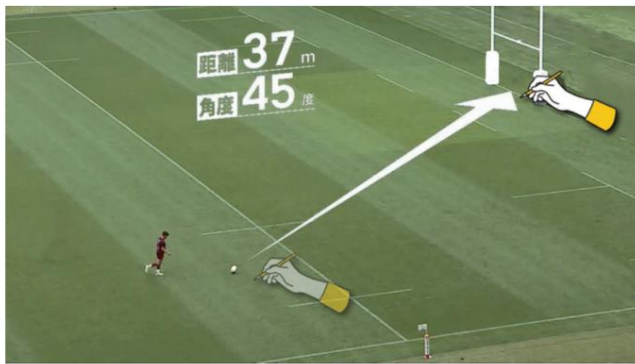


Figure 4. Distance and angle CG

JAPAN/VENEZUELA														
2	TOGASHI Makoto	3	1/4	5	0/0	25	0	0/0	1	0	1	2	2/2	
5	KAWAMURA Yuta	10	2/4	3	2/5	67	0	0/0	3	4	1	3	0/0	
6	BABA Yuki	6	3/5	2	3/5	70	0	0/0	3	4	1	4	0/0	
12	NISHIDA Keisei	9	4/7	4	2/3	33	100	6/6	4	8	5	0	1	0/0
18	INOUE Shuta	3	3/6	5	1/3	0	100	4/4	8	4	1	3	0/0	
19	HIEJIMA Makoto	0	0/1	6	0/0	0	0	0/0	4	2	1	2	0/0	
24	HARA Hitorake	0	0/1	7	2/4	0	0	0/0	2	1	1	1	0/0	
30	WATANABE Kota	2	0/0	8	1/3	0	50	2/4	1	1	1	1	0/0	
31	INUJIMA Arata	0	0/0	9	2/2	0	0	0/0	1	3	1	1	0/0	

日本 ベネズエラ			
36	得点	41	
35.0%	2P	57.9%	
7/20		11/19	
27.8%	3P	31.6%	
5/18		6/19	
87.5%	フリースロー	50.0%	
7/8		1/2	
22	リバウンド	26	
8	ターンオーバー	10	

Figure 5. Subtitle display of basketball score data

By simply touching three points on the intersection of the grid, the field is recognized and internally converted into a plan view. After that, by touching any two points on the field, the system calculates the distance and angle. Not only that, but because the system is able to recognize the field, it can also display CG that follows the field, as shown in Figure 4.

This makes it possible to estimate the distance and angle to the goal from the image of a single camera, which previously required a lot of preparation for distance and angle CG, leading to significant labor savings.

3.2 Surrounding Function

In addition to the function of drawing CG wherever a user touches the screen, we have added a function to automatically recognize textual information within an enclosed area and display subtitles or read the text aloud. These examples are explained below.

3.2.1 Automated Score Data Entry This

function was used to send out subtitles during the 2023 Basketball World Cup. During basketball broadcasts, a huge amount of data, such as each player's scoring rate and three-point success rate, must be selected and reflected in the subtitles.

In the conventional operation, this data was read manually, and the necessary data was selected and reflected in the subtitles. However, there was a limit to how much data could be processed manually.

In some cases, it was not possible to provide the necessary information immediately.

To address this issue, we added a function that recognizes the text information of areas circled with the video pen, thereby speeding up data sharing. First, on-site score data is shared on the screen and entered into the video pen system. Then, by selecting an arbitrary data range by enclosing it, character recognition allows the data to be instantly reflected in the subtitles (Figure 5). This allows anyone to easily reflect data in the subtitles in about one second, contributing to improving the informational value of the content.

Another major strength of this function is its ease of use, as all that is required is to share the data screen and circle it. Normally, to share data at this speed, it is necessary to contact the event organizers in advance and have them allocate a system for data output. However, the larger the event, the greater the costs and operational risks for the event organizers, so this is not the case for international events.

However, by using this method, the entire process can be completed within the company's own system, overcoming this concern.

Furthermore, the ability to select desired data while looking at the screen also enhances the ease of use of this function. Previous attempts have been made to acquire data using image recognition¹⁾, but in order to select a range, it was necessary to look up the x and y coordinates on the screen and enter those values into a program file. This made it difficult to quickly respond to sudden layout changes, but with this function, the task can be completed simply by surrounding the required data while looking at the screen. Taking advantage of this strength, it was possible to flexibly respond to layout changes on the data screen on the day of operation at the 2023 Basketball World Cup.

3.2.2 Automation of Ekiden Course Guides At

the 100th Hakone Ekiden race, held in January 2024, an unprecedented production support system was implemented by adding an automatic voice reading function in addition to the character recognition with free range selection described above. At

the Hakone Ekiden, radio waves from the broadcast van are received at approximately 20 relay points. At each point, broadcast staff are waiting, pointing their antennas at the broadcast van to receive the radio waves. If the position of the broadcast van is lost, there is a risk that the radio waves will be cut off and the broadcast will stop. For this reason, the position information of the broadcast van must be communicated to the broadcast staff one by one, but this information is transmitted using mobile devices, etc., when working at high altitudes. In this case, there is a risk of an accident due to the device being dropped.

Therefore, the current location of the live broadcast information is communicated to the on-site staff as audio information (course guide) via

an intercom. However, in the previous operation, the course guide was read out manually, which required the reading staff to work for a total of approximately 70 hours, including test and rehearsal days. There were also instances where the reading did not proceed smoothly due to mistakes made when the staff lost concentration, and conflicts in the reading timing of the four mobile broadcast vehicles. In order to improve the efficiency of this operation,

a new function was developed that combines a function to recognize the text information in the enclosed part with automatic reading.

[Copyrights to Machine-Translated Content]

The copyright of the original papers published on this website belongs to the Institute of Image Information and Television Engineers. Unauthorized use of original papers or translated papers is prohibited. Please be sure to cite the original publication when referencing.

For details, please refer to the copyright regulations of the Institute of Image Information and Television Engineers.

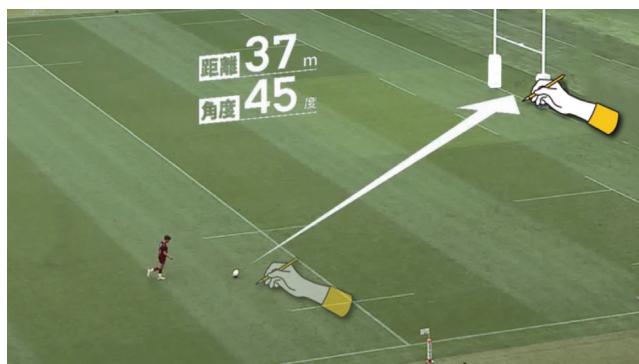


図4 距離角度CG

の交点を3点タッチするだけで、フィールドを認識し、内部で平面図へと変換する。その後、フィールド上の任意の2点をタッチすることで、システム内で距離と角度を算出する。それだけにとどまらず、フィールドを認識できているため、図4のようにフィールドに沿うようなCGも表示できる。

これにより、従来準備が大変であった距離角度CGにおいて、カメラ1台の映像からゴールまでの距離と角度を推定することができ、大幅な省力化につながった。

3.2 囲む機能

タッチした箇所にCGを描画する機能に加え、囲んだ範囲の文字情報を自動的に認識してテロップ表示や音声読み上げなどを行う機能を追加した。以下ではこれらの事例について説明する。

3.2.1 得点データ入力自動化

2023年のバスケットボールW杯において、本機能を活用したテロップ送出を行った。バスケットボール中継では、選手ごとに得点率や3ポイント成功率など、非常に多くのデータを取捨選択してテロップに反映させなければならない。

従来の運用ではこれらのデータを人力で読み取り、必要なデータを取捨選択してテロップへ反映させていた。しかし、膨大なデータ数を人力で処理するには限界があり、出した情報をすぐさま出すことができないケースもあった。

そこで、ビデオペンで囲んだ部分の文字情報を認識する機能を追加し、データ連携の高速化を図った。まず、現地得点データを画面共有し、ビデオペンのシステムに入力する。その後、任意のデータ範囲を囲んで選択することで、文字認識により瞬時にテロップへデータを反映することができる(図5)。これにより、約1秒の速さで誰でも簡単にデータをテロップに反映でき、コンテンツの情報的な価値向上に貢献した。

さらに、必要な作業はデータ画面の共有と囲む操作のみという手軽さも本機能の大きな強みである。通常、このスピードでのデータ連携を行う場合、あらかじめ大会側とコンタクトを取り、データ出力の系統を割いてもらわなければならない。しかし、大会の規模が大きくなるほどコストや大会側の運用リスクが大きくなるため、国際大会などで

JAPAN/VENEZUELA														
2	TOGASHI Makoto	3	1/4	5	0/0	25	0	0/0	1	0	1	2	2/2	
5	KAWAMURA Yuta	10	2/4	3	2/5	67	0	0/0	3	4	0	1	3	0/0
6	BABA Yuki	6	3/5	2	3/5	70	0	0/0	3	4	0	4	0/0	
12	NISHIDA Keisei	9	4/7	4	2/3	33	100	6/6	4	8	5	0	1	0/0
18	INOUE Shuta	3	3/6	5	1/3	0	100	4/4	8	4	0	0	0/0	
19	HIEJIMA Makoto	0	0/1	6	0/0	0	0	0/0	4	2	0	0	0/0	
24	HARA Hitorake	0	0/1	7	2/4	0	0	0/0	2	1	0	0	0/0	
30	WATANABE Kota	2	0/0	8	1/3	0	50	2/4	1	1	0	0	0/0	
31	INUJIMA Arata	0	0/0	9	2/2	0	0/0	1	3	0	0	0	0/0	
										日本 ベネズエラ				
										36	得点	41		
										35.0%		57.9%		
										7/20	2P	11/19		
										27.8%	3P	31.6%		
										5/18		6/19		
										87.5%	フリースロー	50.0%		
										7/8		1/2		
										22	リバウンド	26		
										8	ターンオーバー	10		

図5 バスケットボール得点データのテロップ表示

はその方法での連携が難しい。本手法を用いれば、自社内のシステムだけで完結させることができるため、その懸念を克服できる。

そして、「画面を見たまま所望のデータを選択できる」という点も、本機能の使い勝手の良さを加速させた。従来も画像認識を用いたデータ取得の試みがなされてきた¹⁾が、範囲選択のためには画面内のx, y座標を調べ、プログラムファイルにそれらの値を入力しなければならなかった。そのため、急なレイアウト変更には迅速に対応できない仕様であったが、本機能では画面を見たまま必要なデータ部分を囲むだけで作業が完了する。この強みを活かし、2023年バスケットボールW杯での運用においても、当日のデータ画面におけるレイアウト変更にも柔軟に対応することができた。

3.2.2 駅伝コースガイドの自動化

2024年1月に行われた、第100回大会となる箱根駅伝では、上記のような自由範囲選択による文字認識に加え、自動音声読み上げの機能を付与することでこれまでにない制作支援を試みた。

箱根駅伝では、中継車からの電波を約20ヵ所もの中継ポイントで受信している。各ポイントには、アンテナを中継車に向け電波を受信する中継スタッフが待機しており、万が一中継車の位置を見失ってしまうと電波が遮断され放送が止まってしまう恐れがある。そのため、中継車の位置情報は逐一中継スタッフに伝える必要があるが、高所における作業でモバイル端末等を使用した位置情報の伝達を行う場合、端末の落下による事故のリスクがある。よって、中継情報の現在位置はインカムを通した音声情報(コースガイド)として現場スタッフへ伝えられている。

しかし、従来の運用ではコースガイドを人力で読み上げており、読み上げスタッフにはテストやリハーサル日を含めると合計約70時間もの拘束時間が発生していた。また、集中切れによる読み上げミスの発生や、計4台の移動中継車における読み上げタイミングのバッティングにより、読み上げがスムーズに進行しない場面もあった。

この運用を効率化すべく、囲んだ部分の文字情報を認識する機能と自動読み上げとを組み合わせた新機能を開発し



Figure 6 Automatic reading of course guide for the Hakone Ekiden

First, as shown in Figure 6, the distance display area on the PC screen displaying the distance information of the mobile broadcast van is circled with a video pen. This calls up the place name data linked to the distance information, and then an automated voice reads out the current location.

Additionally, in order to automate the process, several operational innovations were incorporated. First, if the mobile broadcast vehicle remains at the same point due to an athlete falling, for example, the name of the location at that point will continue to be read out, thereby informing the field that there is no system shutdown due to a problem. Furthermore, if the speed of the mobile broadcast vehicle increases, older reading cues will be eliminated and the reading of the most recent location will always be given priority. These innovations

have enabled the automation of course guide reading, eliminating the need for two people to read out loud, which was previously required for testing, rehearsals, and actual offer, performances. The team was able to completely reduce manpower for the first five days. Their strength is their error-free operation even over long periods of time, and they supported stable operation of the live broadcast for a total of 12 hours.

3.3 Application to new program production

This system has been developed mainly to support program production. However, as a new attempt, we have added new functions related to program production. Here we will explain an example of this.

In addition to the Basketball World Cup, the Rugby World Cup in 2023 was also attracting attention. In rugby broadcasts, it was necessary to keep viewers interested even during halftime to prevent them from dropping out. Therefore, an unprecedented attempt was made to implement a segment in which AI determines the most exciting scene in the first half of the match. First, a timer displayed during the match was

circled with a video pen (Figure 7), and the time information was recognized as text. This triggered a speech recognition AI that had previously learned the voices of the spectators to calculate the excitement level during the match in real time every second that passed. The calculation every second was designed to match the competitive nature of rugby, where play is frequently stopped due to fouls, etc. This made it possible to visualize the excitement throughout the entire match and quantitatively evaluate which moments were the most exciting.

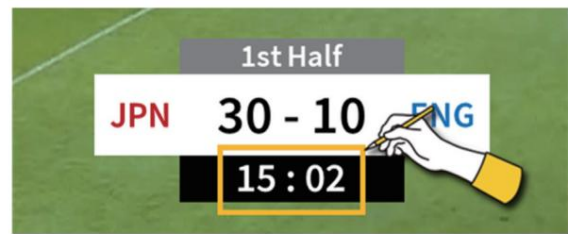


Figure 7 Timer display selection during a match

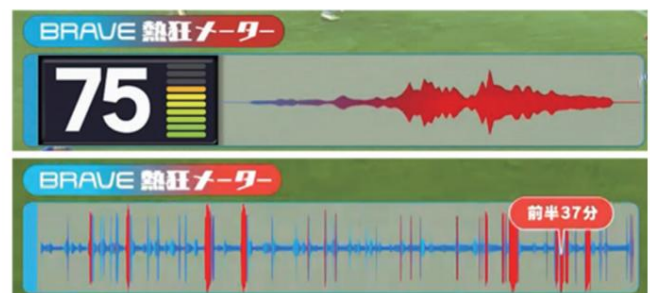


Figure 8. Display of excitement during a

match (Top: Real-time display of excitement figures + subtitles; Bottom: Graphical display of excitement throughout the match)

(Figure 8)

The effects using this function are shown as the "BRAVE Enthusiasm Meter." It was turned into a corner and used on air for all 19 matches. To make it easier for anyone to use, an app with a UI that allows intuitive operation of various functions, such as inputting subtitles and recording history, was developed on just one screen. Taking advantage of the speed and flexibility that only an in-house developer can functions were continually improved throughout the tournament. Thanks to these innovations, it became possible to operate approximately half of all matches without the presence of a technical specialist.

Furthermore, if this production were done manually, someone would have to constantly monitor the match, which would amount to six working days' worth of time for all matches broadcast by Nippon Television. In addition to the new production of quantitative evaluation using AI, the significant contribution to labor savings mentioned above was also an attempt that could only be made with this system.

3.4 Future Developments

As for future developments, we aim to update the monitoring technology by further applying "monitoring excitement during matches." Here,

This section explains its development.

Generally, broadcasting stations have a large number of systems that must be monitored, such as weather information, time signals, and distribution. Most of these are monitored in a master room, but as content becomes more diverse, the number of systems that need monitoring is increasing, and there are limits to how much can be checked manually. Therefore, we are proposing and testing a monitoring system that utilizes a video pen system.

Nippon Television has a specialized channel called "NTV NEWS24" that broadcasts the latest news 24 hours a day.

There is a scroll board as shown in Figure 9, but

[Copyrights to Machine-Translated Content]

The copyright of the original papers published on this website belongs to the Institute of Image Information and Television Engineers. Unauthorized use of original papers or translated papers is prohibited. Please be sure to cite the original publication when referencing. For details, please refer to the copyright regulations of the Institute of Image Information and Television Engineers.

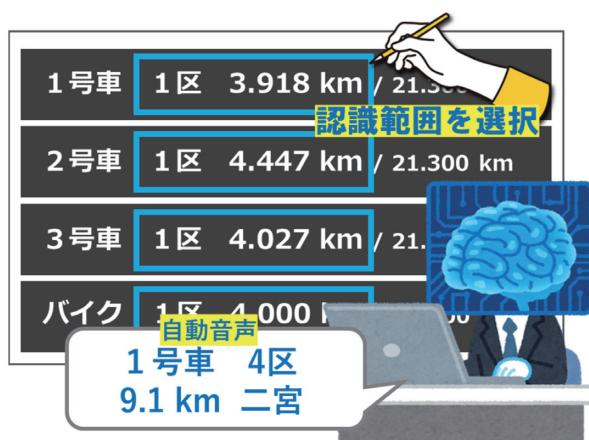


図6 箱根駅伝におけるコースガイド自動読み上げ

た。まず、図6に示すように移動中継車の距離情報を表示しているPC画面の距離表示部分をビデオペンで囲む。これにより、距離情報と紐づけられた地名データが呼び出され、さらに自動音声を用いて現在位置が読み上げられる。

また、自動化を図るため、運用面ではいくつかの工夫を盛り込んだ。まず、選手の転倒等で移動中継車が同じポイントにとどまり続ける場合は同ポイントの地名を読み上げ続けることで、トラブルによるシステム停止が起こっていないことを現場に伝える運用とした。さらに、移動中継車の速度が上昇した場合は、古い読み上げキューを排除して常に最新地点の読み上げを優先的に行う。

これらの工夫によりコースガイド読み上げを自動化し、読み上げに割いていた人員2名×テスト・リハーサル・本番5日間を完全に省人化することができた。長時間でもミスのない運用が強みであり、計12時間にわたる生中継の安定運用を支えた。

3.3 新しい番組演出への応用

主に番組制作支援を軸に展開を続けてきた本システムであるが、新たな試みとして、番組演出に関わる新機能追加を行った。ここではその事例について説明する。

2023年はバスケットボールW杯に加え、ラグビーW杯の注目も高まっていた。ラグビー中継においては、ハーフタイム中も視聴者の興味を引き続け、離脱を防がなければならない。そこで、これまで無かった試みとして「AIが試合前半で最も盛り上がったシーンを決めるコーナー」を実施した。

まず、試合中に表示されるタイマをビデオペンで囲み(図7)、時間情報を文字認識する。この時間をトリガーとして、試合時間が1秒進むごとに、あらかじめ観客の声を学習した音声認識AIが試合中の盛り上がりをリアルタイムで算出する。1秒ごとに算出する仕様としたのは、反則等で頻繁にプレーが止まるラグビーの競技性に合わせたものである。これにより、試合全体を通じた盛り上がりを可視化し、「どの瞬間が盛り上がったか」を定量的に評価でき

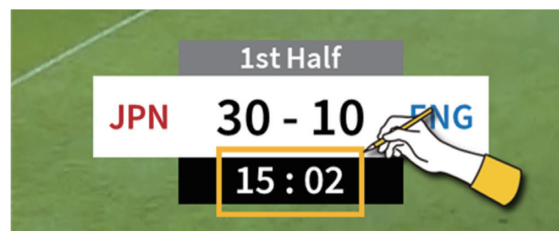


図7 試合中のタイマ表示選択

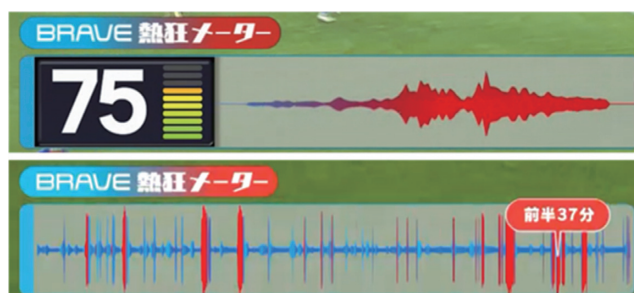


図8 試合の盛り上がり表示

(上：盛り上がりの数値をリアルタイム表示+テロップ表示、下：試合を通しての盛り上がりをグラフ表示)

る(図8)。

本機能を用いた演出は「BRAVE熱狂メーター」としてコーナー化し、全19試合にわたってオンエア使用された。また、誰でも扱いやすくするための工夫として1画面のみでテロップ入力や履歴の記録など、さまざまな機能を直感的に扱えるUIをもつアプリを開発した。本アプリは自社開発ならではのスピード感と柔軟性を活かして、大会期間中も機能改修を重ねていった。これらの工夫により、全試合の約半分以上を技術担当者の立ち合いなしで運用することが可能となった。

さらに、本演出を人力で行うならば、誰かが常時試合を監視し続けていなければならない。日本テレビが放送した全試合で換算すると実働6日分もの拘束時間が生じる。AIによる定量評価という新演出に加え、上記のような省人化に大きく貢献した点も本システムでしかできない試みであった。

3.4 今後の展開

今後の展開として、「試合中の盛り上がり監視」をさらに応用した監視技術のアップデートを目指している。ここでは、その展開について説明する。

一般的に、放送局においては気象情報や時報、配信など監視しなければならないシステムが非常に多い。多くはマスタールーム等で監視が行われているが、コンテンツの多様化に応じて監視が必要なシステムは増えており、人力での確認には限界がある。そこで、ビデオペンシステムを活用した監視システムを提案・テスト中である。

日本テレビは、最新ニュースを24時間配信する専門チャンネル「日テレNEWS24」を有している。このチャンネルには図9に示すスクロールボードが存在するが、「続けて同じ



Figure 9. Automatic monitoring of news distribution

This has contributed to new production effects and has evolved into a comprehensive system that goes beyond production support. Taking advantage of its intuitive interface, it is expected that the range of applications will continue to be flexibly expanded to cover a wide range of content.

〔文 献〕

1) Takayuki Shinoda: "In-house development and operation of AI business support system 'Aidy'",

Journal of Visual and Media Studies, 77, 1, pp.109-113 (2023)

There is a need to monitor whether news is being broadcast or whether any fields are blank.

Therefore, we developed a system that uses a video pen to circle the scrolling area and notifies users of abnormalities with audio information when the above problems occur. The automatic audio notification of the abnormality makes it possible to quickly notice problems even in a master room with a large number of monitors. It is.

In this way, the video pen system can be used to help broadcast stations operate unmanned. In the future, we hope to further expand the monitoring range.

4. Conclusion

Starting with the video pen, various programs were produced under the concept that "anyone can intuitively use new technology."



岸 楓馬 2022年、日本テレビ放送網(株)入社。スポーツCGやリアルタイムCGシステムの運用・管理、新技術開発・導入に従事。



三浦 祐樹 2023年、日本テレビ放送網(株)入社。スポーツCGやリアルタイムCGシステムの運用・管理、新技術開発・導入に従事。



篠田 貴之 2008年、日本テレビ放送網(株)入社以来、番組制作・CG技術部門の現場業務と並行して多数の開発に従事し、約30の社外賞を獲得。正会員。



図9 ニュース配信の自動監視

ニュースが流れていないか」、「空欄になっていないか」を監視したいというニーズがある。

そこで、ビデオペンを用いてスクロール部分を囲むことで、上記トラブルが発生した際に音声情報で異常を知らせるシステムを開発した。自動音声で異常の内容を伝えることができるため、モニタ数が非常に多いマスター室においてもいち早くトラブルに気が付くことができる仕様となっている。

このように、ビデオペンのシステムを活用することで放送局の無人運用にも役立てることができる。今後、さらなる監視範囲の拡大につなげていきたい。

4. むすび

ビデオペンを起点として、「誰でも新技術を直観的に扱うことができる」というコンセプトのもとさまざまな番組

への応用が実現した。新たな演出効果にも貢献し、もはや制作支援だけにとどまらない包括的なシステムへと進化を遂げている。直感的なインターフェースを活かして、今後もあらゆるコンテンツへと柔軟に活用範囲を広げていくことが期待できる。

〔文 献〕

- 1) 篠田貴之：「AI業務支援システム「エイディ」の社内開発と運用」, 映情学誌, 77, 1, pp.109-113 (2023)



岸 楓馬 2022年、日本テレビ放送網(株)入社。スポーツCGやリアルタイムCGシステムの運用・管理、新技術開発・導入に従事。



三浦 祐樹 2023年、日本テレビ放送網(株)入社。スポーツCGやリアルタイムCGシステムの運用・管理、新技術開発・導入に従事。



篠田 貴之 2008年、日本テレビ放送網(株)入社以来、番組制作・CG技術部門の現場業務と並行して多数の開発に従事し、約30の社外賞を獲得。正会員。