

Para Além da Estética, a Presença da Tecnologia Digital nos Programas Televisivos de Transmissão “ao vivo” de Futebol¹

Tatiana Zuardi USHINOHAMA²

Marco ROXO³

Universidade Federal Fluminense, Niterói, RJ

RESUMO

Um marco da tecnologia televisiva na transmissão esportiva “ao vivo” aconteceu na Copa Mundo de Futebol em 2010 quando, desde a captação até a recepção, o processo tecnológico foi digital. Com isso, o público se deslumbrou com uma camada superficial das características dessa tecnologia, a imagem e o som, uma vez que as características profundas dessa tecnologia como a automação, o algoritmo, a ubiquidade e a pervasividade eram invisíveis, pois estavam no processo de produção e distribuição do programa. Com isso, este estudo visa investigar como essa camada mais profunda da tecnologia digital envolveu-se na transmissão televisiva de futebol “ao vivo”. A partir dos dados analisados, verificou-se que essas características invisíveis da tecnologia ao público ocasionaram uma distribuição de uma maior quantidade de informação a respeito do jogo devido a uma apreciação automatizada da partida.

PALAVRAS-CHAVE: Futebol; transmissão televisiva; tecnologia digital; televisão; comunicação.

A Copa do Mundo de 2010 apresentou um show de imagem e som na sua transmissão televisiva, já que os elementos audiovisuais na tecnologia digital eram de melhor qualidade do que os captados e distribuídos pelo sistema tecnologia analógico. Foram 32 câmeras que captaram a imagem do evento em uma proporção de 16:9, sendo que 8 delas possuíam o recurso de super *slowmotion*⁴. E som ambiente do evento esportivo foi capturado por 26 microfones distribuídos em torno do campo. Todos esses equipamentos produziram uma grande quantidade de dados para a equipe televisiva, que precisavam ser nomeados e indexados para imediato ou posterior uso na transmissão.

¹ Trabalho apresentado no GP Comunicação e Esporte, XX Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 43º Congresso Brasileiro de Ciências da Comunicação.

² Doutoranda do Programa de Pós-Graduação em Comunicação da Universidade Federal Fluminense (PPGCOM-UFF), e-mail: tatianazuardi@globo.com.

³ Docente do Programa de Pós-Graduação em Comunicação da Universidade Federal Fluminense (PPGCOM-UFF), e-mail: marcos-roxo@uol.com.br

⁴ O olho humano, parâmetro de comparação, tem uma capacidade de percepção média de no mínimo de 24 quadros por segundo e no máximo de 200 quadros por segundo. A câmera em super *slowmotion* grava as ações dos jogadores em uma quantidade de quadros por segundo que vai de 150 a 1000, o que permite exibir a ação capturada em uma velocidade reduzida durante o *replay* de modo que haja uma excelente descrição estética dos detalhes do movimento.

Desse modo esses dados precisavam ser transformados em informações para operados e aproveitados. Com isso, a tecnologia digital provocou um ponto crítico para os produtores o de como gerenciar todos os dados captados durante os noventa minutos do jogo e durante os 64 jogos do campeonato e transformá-los em informação para ser distribuída rapidamente.

Uma boa gestão desses dados proporcionaria novas opções para a produção dos programas esportivos, promoção das instituições e interação entre público de modo a evidenciar o evento. Uma má gestão poderia, ao contrário, levar a um acúmulo e ao desperdício de material o que desorientaria a equipe de produção, impossibilitando-a de usufruir dessa maior quantidade de dados disponibilizados pelo aumento no número de equipamentos de captura e dos recursos da transformação dos sinais analógicos, visuais e sonoros, em uma representação numérica.

Assim, o objetivo deste estudo é investigar como a camada mais profunda da tecnologia digital, a automação, o algoritmo, a ubiquidade e a pervasividade, envolveu-se na transmissão esportiva “ao vivo a fim de considerar se a aparição desses algoritmos afeta a transmissão e o seu processo televisivo. Para isso, realizou-se um levantamento bibliográfico de pesquisas científicas no banco de dados da CAPES que abordavam a temática a partir da combinação (e/e/e) das palavras-chave: Broadcasting, Sport, Digital, com uma delimitação de temporal até o ano de 2018, ano da última Copa do Mundo. Foram encontrados 32 artigos abordando a temática em diversos esportes (futebol (13), esporte em geral (6), basquete (5), beisebol (3), atletismo (1), críquete (1), hockey (1), tênis (1), vela (1)). Em virtude disso, houve a necessidade de se aplicar um segundo filtro, selecionando apenas os artigos direcionados à modalidade futebol. Por isso, o corpus da pesquisa constitui-se de 13 artigos, entre os anos de 2003 e 2018, que investigaram a camada profunda da tecnologia digital na transmissão televisiva “ao vivo” de futebol.

Todos os artigos encontrados pertencem a área de conhecimento da Ciência da Computação, pois visam pesquisar uma estrutura de extração de informações a partir do vídeo. Não foi encontrado neste levantamento pesquisas envolvendo essa temática na área da Ciências Humanas aplicadas, Comunicação, que objetivam discutir a inserção dessa ferramenta na sociedade da perspectiva dos processos comunicacionais.

Para desenvolver o estudo neste texto, propõe-se que, na primeira parte, o texto disserte sobre a relação entre a tecnologia digital, a especificidade do programa e as características do futebol; na segunda parte, explica-se como o processo de análise de

vídeo ocorre para o futebol. O capítulo termina por discutir, sob a perspectiva da comunicação, a presença dessa estrutura tecnológica digital aplicada ao programa de transmissão de futebol “ao vivo”, e o mesmo questiona: quem está determinando os parâmetros relevantes da automação de um processo de análise de vídeo na produção de um algoritmo que busca interpretar um programa televisivo a fim de produzir informações extras, o produtor cultural de comunicação ou os programadores do sistema de engenharia computacional e da matemática?

1 - TECNOLOGIA DIGITAL, PROGRAMA TELEVISIVO E ESPORTE

A tecnologia digital na transmissão televisiva do futebol “ao vivo” proporcionou ao público, a partir da Copa do Mundo de 2010, uma mudança estética⁵ na forma visível de assistir ao jogo, uma distribuição de mais informações sobre os jogos⁶ e, também, vídeos com os melhores momentos, em um curto período de tempo após o encerramento das partidas, pela instituição organizadora (FIFA).

Nessa Copa, a FIFA dispôs essas informações adicionais em uma plataforma na internet (site) que distribuía uma ampla variedade de conteúdo e propiciava uma interatividade em concomitância com o evento. Esse sistema de comunicação exibia informações relacionadas aos jogos do campeonato, fornecendo detalhes estatísticos e visuais coletados a partir das imagens captadas pelas câmeras instaladas no estádio. Eram desde informações simples de serem apuradas como o número de chutes em direção ao gol de cada time, as faltas, a posse de bola, os impedimentos, os escanteios, os cartões, até um levantamento de quantitativo de desempenho técnico de times e jogadores que manualmente é laborioso de ser computado como o registro de ações do jogo em ordem temporal, o número total de passes de cada jogador, o percentual de passes corretos de cada equipe, a distância percorrida, as velocidades máxima e média de cada jogador na partida e o tempo de posse de bola de cada equipe.

Essas informações estatísticas dos jogos eram divulgadas pela FIFA rapidamente, como consequência de um processo automatizado de identificação, indexação, seleção, quantificação, separação e armazenamento estabelecido por um algoritmo executado durante os jogos. Esse procedimento automatizado de análise e extração de informações

⁵ USHINOHAMA, T. (2015), Narrativa Audiovisual da Transmissão Televisiva de Futebol. Saarbrücken: Novas Edições Acadêmicas.

⁶ Disponível em: <<https://www.fifa.com/worldcup/news/fifa-com-attracts-over-quarter-billion-visits-the-world-engages-online-1273696>>. Acesso em: 08 set. 2020.

pôde ser executado utilizando e combinando diversos parâmetros dos dados disponíveis no vídeo, de modo a entregar um resultado (conhecimento) em três temporalidades diferentes: simultaneamente ao jogo ou após o seu término (em curto ou longo intervalo de tempo).

Para que isso acontecesse, houve uma diversidade de procedimentos que precisaram ser programados a fim de que vincular uma direta efetividade da tecnologia digital à análise do vídeo, em busca de transformar os dados no vídeo (imagem, som, metadados) em informações do jogo. Essa estrutura requeria que o algoritmo combinasse os dados para identificar as unidades semânticas do jogo em um elo entre as pessoas e o espaço, entre elas e o tempo, entre elas e a bola, e entre elas durante a partida de maneira que essas relações propiciem extrair informações que podem ser relevantes para o interesse público em geral. A apuração dessa análise automatizada resulta na detecção de ações relevante do jogo, apesar das limitações técnicas ainda presentes em um procedimento automatizado, mas que já precisa alcançar uma alta taxa de confiabilidade.



Diagrama 1 – Estrutura de um sistema de análise do vídeo destinado ao futebol.

Fonte: elaboração própria (2020).

Por isso, para traçar um conjunto de regras e procedimentos lógicos para interpretar os dados do vídeo, os engenheiros da computação/processamento de dados precisaram entender o raciocínio estabelecido pelo programa de transmissão esportiva “ao vivo” que capturava os vídeos e o futebol, destacando suas características gerais, antes de definir algoritmos confiáveis para a interpretação do jogo.

Assim, eles apuraram que o futebol possuía uma dinâmica de jogo, com ações em fluxo contínuo, realizada por pelo menos, 25 elementos em campo, ou seja, o ato esportivo de um time não tem um início e um fim bem delimitados ou breves, de modo que as ações ocorrem constantemente, em todo o campo - diferentemente do basquete, em que, após uma cesta ou um arremesso, um ato esportivo termina e outro é iniciado, acontecendo dentro de uma temporalidade média de 24 segundos. Diante disso, a prática do futebol dificulta ao processo automatizado identificar o início e o fim de atos

esportivos significativos, pois o sistema requer pontos de referência claros para estabelecer uma interpretação semântica padrão para o jogo.

Um fator que equilibra essa dificuldade de interpretação do sistema é a regra do jogo, que é fácil, já que, em um espaço delimitado por diversas linhas balizadores, a equipe vencedora é aquela que fizer o maior número de gols dentro de uma temporalidade de 2 períodos de 45 minutos. Essa regulamentação acaba instituindo como elemento fundamental a bola, que se encontra em constante movimento e pode ser rastreada pelo campo, em função da sua cor. A presença da cor permite, também, identificar e rastrear os jogadores e o árbitro pelos uniformes, e destacar o número estampado nas costas dos atletas nas imagens.

Portanto, na medida em que as ações em torno da bola vão se desenvolvem pelo campo, a transmissão televisiva “ao vivo” vai traduzindo o jogo, em imagem e som, sem interferir na performance dos atletas. Essa convenção faz com que esse programa televisivo de entretenimento demande um imediato consumo do público, pois o seu principal atrativo é apresentar a construção do resultado simultaneamente à sua realização devido a um encadeamento das ações dos times.

Essa configuração faz com que um dos principais recursos dessa mediação seja propiciar que ações relevantes possam ser reexaminadas por meio do replay, em diversos ângulos. O que faz do replay um dos primeiros pontos de referência a ser utilizado pelo o processo de automação para identificar um ato semântico significativo no jogo, uma vez que rever uma ação esportiva indica importância para o jogo.

Com o término do jogo, o resultado é definido e o programa televisivo da transmissão esportiva “ao vivo” se encerra e perde seu valor para o público, que passa a buscar por informações sobre a atuação dos jogadores e as ações que resumem o jogo, os melhores momentos. Essa procura por uma síntese do evento esportivo (um conhecimento sobre a partida) foi um dos principais motivos que levaram os acadêmicos e o mercado a pesquisar por procedimentos que automatizassem e agilizassem a produção de um resumo dos eventos esportivos, uma vez que a quantidade de eventos tem aumentado cada vez mais e os dispositivos individuais surgem como uma possibilidade difundir essas informações especializadas e personalizadas para o fã do esporte.

Afinal, a televisão é um meio de comunicação de massa que, pela sua concepção, acaba generalizando e limitando a distribuição da informação, muitas vezes, conforme a

proposta dos canais (generalista ou segmentada), a grade de programação, ao tempo dos programas e compreensão do seu público.

2 - O PROCESSO DE ANÁLISE DE VÍDEO PARA O FUTEBOL

Desde 2003, as pesquisas acadêmicas internacionais buscam uma configuração consistente e precisa para automatizar uma análise de vídeo do futebol, a fim de extrair inúmeras informações sobre o jogo (aspectos técnicos dos jogadores e resumo do jogo) e até conhecimentos (análises táticas e de desempenho dos jogadores). O ponto de partida da detecção automática de sequências contendo as partes interessantes do jogo depende dos parâmetros estabelecidos e de seu entrelaçamento estabelecido pela programação (algoritmos) que interpretará os elementos do vídeo (dados) e atribuirá um significado (ponto de vista semântico). Essa operação tem sido aprimorada, principalmente em função do aumento do poder dos computadores nas últimas décadas, que processam cada vez mais tipos de elementos no vídeo, simultaneamente, propiciando que os objetos semânticos consigam ser extraídos corretamente e o conteúdo da cena possa ser reconstruído.

Para entender a influência desse processamento automatizado de informações a partir do vídeo no programa televisivo da transmissão de futebol “ao vivo”, é preciso entender um pouco quais os dados de vídeo que serão manipulados, sua origem, ferramentas utilizadas e o resultado da aplicação desse sistema. Por isso, nessa seção, busca-se explicar como o processo de automação da análise de vídeo ocorre no futebol a partir das informações propiciadas pelas pesquisas científicas reunidas no corpus desse estudo, as quais ilustraram, com exemplos, as etapas do processo.

Os pesquisadores têm proposto duas formas de coletar os dados para realizar o processo da análise de vídeo no esporte. A primeira forma é a utilização do vídeo produzido pela própria transmissão televisiva “ao vivo” do jogo. E a segunda é mediante a instalação de câmeras fixas, de alta resolução, em torno do campo, para capturarem os movimentos dos participantes do jogo (bola, jogadores, árbitros)⁷, de modo que esses vídeos se concentrem em um sistema de processamento.

Por se tratar de uma captura de ações esportivas por meio de uma aparelhagem mecânica de imagem e som, ambos os procedimentos estão sujeitos às limitações técnicas

⁷ Essa opção apresenta uma robustez na coleta de dados propiciando aos produtores terem os dados transformados em informações em tempo real, entretanto, há um alto custo financeiro a ser disponibilizado por esse sistema.

em função das diferentes condições de iluminação no ambiente, oclusão de um elemento por outro, a rápida ocorrência de eventos, tamanho pequeno dos objetos no vídeo, deformidade na captura das imagens pela câmera. Todas essas adversidades podem ser compensadas pela supervisão de um ser humano, transformando o processo em semiautomatizado.

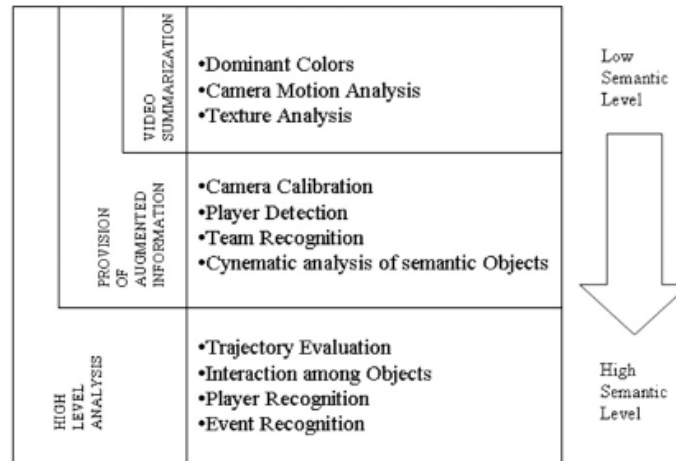


Figura 01 – Esquema de níveis de interpretação semântica do vídeo. Fonte: D’ORAZIO; LEO (2010)

No entanto, há o empenho dos pesquisadores e do mercado privado (times, jogadores, empresas de comunicação) em produzir um processo automatizado, pleno e fidedigno, de análises de vídeo digital do futebol, que ocasione na possibilidade de se extrair diferentes tipos de informações do jogo, em uma hierarquia crescente de complexidade, que pode estar associado⁸ ou não⁹ a um programa televisivo “ao vivo”. Para cada nível de interpretação semântica dos dados, uma estrutura de processamento de análise pode ser aplicada ao vídeo. Os pesquisadores estabeleceram três níveis, iniciando-se no baixo nível semântico e terminando no alto nível semântico (figura 1). A hierarquia dos níveis semânticos tem relação direta com a complexidade na programação de um algoritmo, que estabelecerá um entrelaçamento simultâneo de análises de dados, visando à identificação de um ou múltiplos elementos para a interpretação semântica do jogo (trajetória de um jogador específico, se é ataque ou defesa; ou da bola, se é chute ou lançamento; ou de todos os jogadores, em uma disposição tática de ataque contra defesa). Quanto maior for a necessidade de precisão e segurança em reconhecer um elemento semântico do jogo (pênalti, gol, sistema tático) e extraí-lo, mais complexa será a programação e a maiores serão as exigências de recursos computacionais.

⁸ Exemplo: resumo da partida, melhores momentos do jogo, da equipe ou do jogador, análise tática do jogo ou dos times.

⁹ Exemplo: Estatísticas do jogo, análise técnica do time ou de cada jogador.

O primeiro nível de análise de vídeo, estabelecido para um processo de automação de detecção de eventos, é feito a partir da identificação de elementos da linguagem audiovisual, estabelecida pela transmissão televisiva de futebol, como a presença de *replay*, planos, movimento de câmera, locução e gerador de caracteres, desconsiderando a presença de objetos semânticos no vídeo e uma interação entre eles.

Assim, a partir da transmissão televisiva, Li et al. (2003) se propuseram a interpretar os dados no vídeo com base nas marcas produzidas pela linguagem audiovisual, estabelecida por uma transmissão de futebol “ao vivo”, como: a presença do logo em um efeito de transição um entre as imagens de tempo real e o *replay*; a presença de um enquadramento fechado ou de uma sequência deles a partir do historiograma de cor do frame; e a frequência das ondas sonoras produzidas durante uma locução; a fim de detectar momentos em que houve uma interrupção no jogo de ações contínuo. Os pesquisadores identificaram que esses elementos na transmissão televisiva são índices de pausas. E, nesses momentos podem ter ocorrido ações relevantes na partida a serem destacadas.

Por exemplo, a partir do historiograma das cores informado a cada frame, os pesquisadores verificaram que era possível determinar o tipo de plano (abertos ou fechados) e o objeto principal enquadrado (campo ou pessoas) em uma transmissão televisiva. Essas informações são descritas no metadados do frame e enviadas junto à compressão de dados do vídeo para serem remontadas quando o arquivo de vídeo for lido por um sistema computacional de descompressão. Por isso, a partir da avaliação desse dado, se identifica o fluxo do andamento do jogo.

Isso propicia inferir que os planos abertos da transmissão possuem grande quantidade do campo sendo enquadrada, significando que a bola está em jogo. Quando a bola está parada, os planos são fechados e os objetos enquadrados são pessoas, com quantidade pequena do campo no frame. Assim, uma imagem aberta na transmissão de futebol possui uma alta quantidade de pixel verde. Já, um enquadramento fechado possui uma quantidade pequena de pixels verdes. Mas, e os planos intermediários, como o Plano Conjunto, que podem ter tanto a bola em movimento quanto a bola parada, como mensurar?

O estudo de Ekin et al. (2003) aprimorou essa identificação dos planos da transmissão a partir da detecção da cor dominante na faixa horizontal da região central no frame, uma vez que o Plano Conjunto (- verde) estava sendo utilizado pela transmissão,

tanto em momentos de bola em movimento quanto de bola parada. Com esse ajuste, o Plano Conjunto seguido de um plano aberto (+ verde), a bola estava rolando. E Planos Conjuntos, seguidos de planos fechados ou enquadramentos da torcida (-verde), indicam que a bola estava parada. Outro aspecto aperfeiçoado foi a identificação do *replay*, com o método de verificar um acentuado no número de frames repetidos, uma vez que, para explicitar uma ação, se reduz a velocidade do vídeo, com um aumento na repetição das imagens.

O uso desses componentes do vídeo¹⁰ é, portanto, denominado parâmetros de baixo nível semântico, pois é possível extrair um resumo das ações do jogo e os melhores momentos, de forma automática, entrelaçando os elementos estabelecidos por uma transcrição do evento, a linguagem audiovisual da transmissão televisiva, para obter um resultado preciso após o término do jogo.

Já o segundo nível de processamento, busca extrair e/ou expor informações extras sobre o jogo em função da identificação dos objetos ou local pontuais no vídeo que possuam significado no jogo e da coleta desses dados específicos. Esse processo de elaboração das informações extras requer uma programação mais complexa, uma vez que é necessário identificar, reconhecer e interpretar um determinado objeto semântico da cena para se produzir respostas adequadas que são desde a exibição das informações estatísticas (gol, chutes, vitórias), numéricas (velocidade, distância) até o esclarecimento de lances de impedimento. A presença dessas informações na transmissão de televisão especializa¹¹ o programa nos dois sentidos da palavra, tornar-se especialista em algo e o particulariza, pois apresenta um levantamento técnico do jogador, quando ele for reconhecido na imagem e acrescentam referências técnicas específicas do jogo que complementam os seus serviços de comunicação, durante o programa (no intervalo da partida ou após seu encerramento).

A identificação dos objetos importantes do futebol no vídeo como a bola, os árbitros e os jogadores pode ser feita por meio das cores ou da representação visual (número, formato, desenho). Entretanto, para promover a detecção de uma representação visual, o reconhecimento e o rastreamento da sua trajetória de movimento (a bola e os

¹⁰ Os recursos considerados para o processo de inferência são região de cor dominante, movimento de câmera, intensidade da textura, logotipo, linhas paralelas, placar na imagem, objetos pretos, modificação na frequência no áudio da locução e cena estática ou repetida.

¹¹ Exemplo de informações extras exibidas na transmissão que complementam a transmissão: as distâncias entre o jogador e o gol em um lance de falta; velocidade da bola em um chute no gol ou lançamento; rever a linha de impedimento em uma jogada.

jogadores), faz-se necessário que haja uma calibração entre as imagens captadas pela câmera e o campo, pois, devido à propriedade de convergência da luz nas lentes das câmeras, a imagem sofre uma distorção de perspectiva em relação ao campo (BATTIKH; JABRI, 2011).

Após a calibragem das câmeras e entre as câmeras, os elementos identificados na imagem são transformados em representação gráfica para serem mensurados e depois retornam ao vídeo, convertidos em formas visuais a fim de coincidir com a visão da câmera (o posicionamento da câmera, a orientação e o comprimento focal). Assim, essas informações extras, para serem disponibilizadas na transmissão, requisitam, portanto, que a câmera esteja fixa e não apresente recursos variados de pan, *tilt*, zoom ou baixa resolução de pixels, para que não dificultem a sua aplicação sobre a imagem.

Outro mecanismo desenvolvido neste nível de análise do vídeo foi a detecção das zonas do campo nas imagens televisivas, que podem ser realizada pela disposição das linhas do campo enquadradas pela câmera central (BERTINI et al., 2005). Já Ekin et al. (2003) propõem identificar a cobrança de pênalti com a detecção das três linhas paralelas no enquadramento, delimitadoras do espaço no jogo (linha de fundo, linha da pequena área, linha da grande área), uma vez que a imagem capturada por uma câmera central estabelece uma razão entre os tamanhos, a distância e o paralelismo dessas linhas, que não se equipara a nenhuma outra relação entre linhas do campo. E outra forma de reconhecimento das áreas do campo também pode ser realizada pela identificação dos textos nas placas de publicidade em torno do campo (BALLAN et al., 2010).

Já a detecção dos jogadores ou do árbitro pode ser estabelecida pela distinção entre os uniformes (EKIN et al., 2003) e/ou a presença dos números nos jogadores (BALLAN et al., 2010).

O terceiro nível de processamento dos dados no vídeo destaca-se por realizar uma interpretação considerada de alto nível semântico, uma vez que se buscam, por meio da análise dos vídeos, respostas sobre a atuação dos times e/ou jogadores, ou seja, uma análise tática. Para que isso ocorra, primeiro é necessário a detecção exata e o rastreamento preciso da bola e, também, de cada jogador em campo. Em seguida é possível depreender uma análise complexa a partir do processamento dos dados originados dos jogadores e equipes, visando informar os técnicos, os programas de televisão e, atualmente, inclusive, a arbitragem sobre detalhes de lances, jogadas e irregularidades na dinâmica do jogo.

Com esse nível de processamento dos dados de vídeo torna-se possível auxiliar a arbitragem em mais situações¹², por meio da computação, uma vez que o futebol é um jogo repleto de ações interpretativas em que, cada vez mais, situações duvidosas aparecem nas partidas, de modo que as mesmas podem ser esclarecidas pelas imagens. Por exemplo, quando um jogador obstrui a trajetória da bola adversária com a parte superior do corpo durante uma jogada na sua área de defesa. Foi com a mão? É pênalti? Com a implantação desse nível no sistema de processamento de análise de vídeo é possível realizar o rastreamento da bola e do jogador e, também, fornece uma análise precisa da postura do jogador e da sua interação com a bola.

Entretanto, os estudos (OSKOUIE et al., 2012) têm mostrado que “(...) as imagens de transmissão nem sempre podem ser utilizadas por esses sistemas porque a posição das câmeras não garante a avaliação exata da posição do jogador no campo de jogo”¹³, devido a problemas específicos de identificação de regiões de cor correspondentes ao elemento marcado; estimativa do ângulo e inclinação da câmera da transmissão e os parâmetros de zoom; mudanças de iluminação do ambiente; os movimentos lentos e rápidos dos jogadores e/ou com a bola que podem borrar na imagem da câmera, dificultando a leitura correta da cena ou, ainda, a troca entre as câmeras e/ou planos pode fazer o sistema embaralhar-se e não reconhecer precisamente os elementos marcados do jogo.

Já em relação às interações da bola com os jogadores, devem-se considerar problemas em que a interação entre eles possa causar oclusões e sobreposições de um elemento da imagem em relação a outro, pois a bola representa um pequeno número de pixels na imagem, o que dificulta a detecção da mesma. Uma proposta que vem sendo aplicada é basear a identificação da bola na evolução da sua trajetória. Um fator limitante nesse método são as câmeras com captura de frame de frequência baixa (-30f/s), pois a bola pode aparecer em um frame e desaparecer no frame seguinte, dificultando o processamento desse dado.

Por isso, duas perspectivas foram impulsionadas. A primeira foi coletar de dados de imagem a partir de câmeras fixas, posicionadas ao redor do campo, que, calibradas e sincronizadas, permitem reconstruir o cenário do jogo virtualmente, em 3D. Isso propicia

¹² Auxílio externo a arbitragem com o sistema de detecção do gol implantado na Copa do Mundo de 2014.

¹³ Tradução livre de: “(...) broadcast images cannot always be used by these systems because the position of the cameras does not guarantee the exact evaluation of the player position in the playing field” (D’ORAZIO; LEO, 2010, p. 2922).

gerar uma representação precisa do movimento e da sua segmentação em ações (SANKOH et al., 2015).

A segunda possibilidade é manualmente delinear, no quadro de cada câmera a região correspondente aos jogadores, ao árbitro e à bola de modo que permita que o sistema aprenda o modelo de referência (SPAGNOLO et al., 2007).

Em ambos os casos, essa informação especializada captada por esses métodos externos à transmissão televisiva podem ser incorporados, futuramente, simultaneamente à produção do programa, aprofundando as características quantitativas do jogo (escaltes dos times e jogadores, distância percorrida pela bola durante o lançamento).



Figura 02 e 03 – Ilustra a identificação manual dos jogadores e árbitro no enquadramento da câmera.
Fonte: SPAGNOLO et al. (2007).

Assim, o sistema computacional processa as informações identificadas no vídeo, interpretando os eventos ou os acontecimentos presentes em um dos 16 tipos pré-delimitados: gol, pênalti, impedimento, escanteio, cartões, lateral, drible, passe, cabeçada, reinício de jogo, substituição de jogador, passe longo, recepção de bola, posse de bola e chute a gol. Em seguida à detecção, há o registro das marcas de data e de hora, e outras informações necessárias dos eventos, que indexam, em marcas, o vídeo para serem recuperados posteriormente.

Esses processos automatizados ou semiautomatizados de análise de vídeo possibilitados pela tecnologia digital, propiciam manusear, agilmente, um grande volume de dados, extraindo diversos níveis de informação e distribuindo-os para o público. Isso se converte em quatro tipos de resultados, ao utilizar uma dessas três formas de análise semântica do vídeo de futebol apresentadas acima.

O primeiro tipo é um videoclipe, que consiste no resumo das ações no jogo. O segundo tipo de conhecimento inclui informações estatísticas sobre eventos ocorridos na partida. O fornecimento de informações táticas relacionadas às estratégias do time e do

jogador é considerado a terceira saída do sistema de análise de vídeo de futebol. O último é a indexação e a recuperação de vídeos de futebol.

3 - DISCUSSÃO

Os produtos concebidos pela aplicação de um processo automatizado de análise do vídeo do futebol espalhem-se pelos meios de comunicação digitais de massa, segmentado ou de nicho, atingindo um público geral ou especializado em futebol em prazos que chegam até a simultaneidade com transmissão do jogo. Isso foi possível “(...) quando as invenções técnicas correspondentes permitiram a transposição das barreiras naturais” (KOSELLECK, 2001, p.148), ou seja, a análise do jogo por um processo automatizado computacional promoveu, em um curto espaço de tempo, um aumento na produtividade de informações a respeito de um evento cerimonial de interesse da sociedade.

Essa aceleração na circulação das informações propiciou, conseqüentemente, uma aceleração nas interações humanas, uma vez que acontecimentos e eventos, belos ou duvidosos, das partidas e da competição relevante se tornassem assunto das conversas presenciais e virtuais nas sociedades. O que, concomitantemente, tornou-se uma forma de mediação que provocou uma nova dimensão de experiência temporal em torno do evento, já que, ao final da partida, foram compartilhados uma descrição do jogo com comentários, um *clip* com os melhores momentos, os dados estatísticos e técnicos, uma análise tática de lances importantes¹⁴ para um público em diversos dispositivos, canais, instituições. Por isso, essa “(...) aceleração é mais que mera mudança, e mais que mero progresso. Ela qualifica o ‘processo da história’” (KOSELLECK, 2001, p.153), transformando as características profundas da tecnologia digital em um marco que incita uma resposta no cenário de mediação comunicacional em torno da cobertura esportiva.

Assim, ao identificar como esse processo automatizado de análise de vídeo envolve-se com a transmissão televisiva, torna-se fundamental inferir reflexões corretas a respeito dos rumos que a comunicação esportiva, principalmente o programa de transmissão esportiva “ao vivo”, irá assumir no futuro a partir da incorporação dos

¹⁴ Mesa tática presente nos programas esportivos no Brasil em 2013. Disponível em: <<https://exame.com/tecnologia/globo-utilizou-novas-tecnologias-durante-copa-das-confederacoes/>>. Acesso em: 17 set. 2020.

recursos da tecnologia digital. Uma resposta pode estar em discutir qual é o denominador comum desses dados.

O processo de automação de análise de vídeo é executado, principalmente, a partir do vídeo produzido pelo programa televisivo da transmissão “ao vivo” do futebol, em que a aplicação dos níveis de análise e “filtros”, seleciona, distingue e extraí objetos relevantes em informações extras a serem distribuída para o público. A dependência desse processo de automação do vídeo da transmissão televisiva do jogo surge já no primeiro nível de análise do vídeo, uma vez que os programadores utilizaram os elementos da linguagem audiovisual como indícios de importância para compactar os noventa (90) minutos do jogo em momentos significantes, de forma rápida, precisa e segura para o público.

Essa leitura dos componentes do vídeo não requer que a máquina realize uma real interpretação semântica do jogo, mesmo no último nível desse processo, apesar das denominações utilizadas para explicar as diferenças de níveis. A função da máquina e da sua programação resume-se em identificar os elementos do jogo, mensurar dados programados, quantificá-los, armazená-los, projetar uma perspectiva e recuperá-los com grande eficiência, ou seja, auxiliar o ser humano a explorar mais intrinsecamente os dados, fornecendo material quantitativo para uma análise especializada humana. Afinal, por enquanto, o sistema computacional não consegue compreender os movimentos humanos, as suas ações esportivas e a configuração tática a fim de interpretá-las para selecionar automaticamente entre as câmeras a melhor imagem para descrever a ação esportiva e, por exemplo, assumir o papel do diretor de imagem em uma transmissão esportiva “ao vivo”.

Assim, por enquanto, o intuito para desenvolver essa tecnologia é que ela auxilie a promover uma cobertura televisiva do evento esportivo comprometida em fornecer o melhor da ocasião, de modo que o telespectador tenha, na medida do possível, acesso ao jogo em andamento e a toda a competição, de forma direta, clara e livre de inferências visuais. Isso faz com que o foco do uso das características profundas da tecnologia digital encontra-se em acelerar o processo de fornecimento de informações cada vez mais específicas sobre o evento e não de automatizar o processo de produção da transmissão “ao vivo”, pois essa tecnologia está, ainda, em fase de implantação.

Essa organização cobertura esportiva televisiva faz da transmissão audiovisual “ao vivo” do futebol a gênese da informação e a nave-mãe da comunicação esportiva, a partir da concepção de Jenkins (2008). E o manuseio das informações processadas, de

forma automatizada a partir da análise do vídeo e distribuídas ao público, podem adquirir diferentes propósitos, o de ser uma complementação da transmissão, quando divulgam as estatísticas do jogo, dos times e dos jogadores. Questionadora de peças do jogo na medida em que reconstrói a ação esportiva detalhando as interações entre os objetos envolvidos (bola e jogadores, jogadores dos times e jogadores e arbitragem). Ou ainda, padronizadora, ao distribuir uma seleção particular dos melhores momentos para o mundo e restringir o acesso ao vídeo original do jogo, direcionando uma memória, apesar da abundância material de informações distribuídas para o público, por uma mesma instituição, a FIFA, a organização esportiva de futebol.

REFERÊNCIAS

- BALLAN, L.; BERTINI, M.; DEL BIMBO, A.; SERRA, G. Semantic annotation of soccer videos by visual instance clustering and spatial/temporal reasoning in ontologies. In: **Multimedia Tools and Applications**, vol. 48, August 2009, pp. 313-337.
- BATTIKH, T.; JABRI, I. Camera calibration using court models for real-time augmenting soccer scenes. In: **Multimedia Tools and Applications**, vol. 51, January 2010, pp. 997-1011.
- BERTINI, M.; CUCCHIARA, R. DEL BIMBO, A.; PRATI, A. An integrated framework for semantic annotation and adaptation. In: **Multimedia Tools and Applications**, vol. 26, August 2005, pp. 345-363.
- D'ORAZIO, T.; LEO, M., A review of vision-based systems for soccer video analysis. In: **Pattern Recognition**, vol. 43, Issue 8, August 2010, pp. 2911-2926.
- EKIN, A.; TEKALP, A.M.; MEHROTRA, R., Automatic soccer video analysis and summarization. In: **IEEE Transactions on Image Processing**, Vol. 12, Issue 7, July 2003, pp. 796-807.
- JENKINS, H. **Cultura da Convergência**. Tradução: Susana Alexandria. 2ªed. São Paulo: Aleph, 2009.
- KOSELLECK, R. **Estratos do Tempo: Estudos sobre a história**, Tradução: Markus Hediger, Rio de Janeiro, Contraponto, 2000.
- LEONARDI, R.; MIGLIORATI, P.; PRANDINI, M. Semantic indexing of soccer audio-visual sequences: a multimodal approach based on controlled Markov chains. In: **IEEE Transactions on Circuits and Systems for Video Technology**, Vol. 14, Issue 5, May 2004, pp. 634-643.
- LI, B.; PAN, H.; SEZAN, I. A general framework for sports video summarization with its application to soccer. In: **IEEE International Conference on Acoustics, Speech, and Signal Processing**, Hong Kong, China, 6-10 April 2003.
- OSKOUIE, P.; ALIPOUR, A.; MOGHADAM, A. Multimodal feature extraction and fusion for semantic mining of soccer video: a survey. In: **Artificial Intelligence Review**, vol. 42, March 2012, pp. 173-210.
- REHMAN, A.; SABA, T. Features extraction for soccer video semantic analysis: current achievements and remaining issues. In: **Artificial Intelligence Review**, vol. 41, March 2012, pp. 451-461.
- SPAGNOLO, P., MAZZEO, P.L., LEO, M., D'ORAZIO, T., Unsupervised algorithms for segmentation and clustering applied to soccer player's classification. In: **Proceeding of the International Conference on Signal Processing and Multimedia Applications**, Barcelona, Spain, 28-31 July 2007, pp. 129-134.
- SANKOH, H.; NAITO, S.; HARADA, M.; SAKATA, T.; MINOH, M. free-viewpoint video synthesis for sport scenes captured with a single moving camera. In: **ITE Transactions on Media Technology and Applications**, vol. 3, Issue 1, 2015, pp. 48-57.
- WAN, K.; XU, C. Efficient multimodal features for automatic soccer highlight generation. In: **Proceedings of the 17th International Conference on Pattern Recognition**, Cambridge, UK, 26-26 Aug. 2004.
- WANG, Z.; YU, J.; HE, Y.; GUAN, T. Affection arousal based highlight extraction for soccer vídeo. In: **Multimedia Tools and Applications**, vol. 73, July 2013, pp. 519-546.