

MPEG IoMT

Lucas Fernandes Martins¹

lucasf.martins94@gmail.com

¹Instituto de Computação – Universidade Federal Fluminense (UFF)
Niterói – RJ – Brasil

***Abstract.** This work aims at describing the main motivations that led the MPEG group to create a standard proposal to investigate the main challenges imposed by the new IoT applications involving multimedia information known as IoMT. Its goals as well as the main existing use cases proposed by the proposal are also addressed. An analysis of the state of the art is done to collect the main studies done in IoMT as well as the results obtained in these works. We also explore the concepts of Mthings and Mwearables defined by the standard proposal that are used in IoMT applications.*

***Resumo.** Este trabalho tem por objetivo descrever as principais motivações que levaram o grupo MPEG a criar uma proposta de padrão para investigar os principais desafios impostos pelas novas aplicações IoT envolvendo informações multimídia conhecido como IoMT. Os seus objetivos bem como os principais casos de uso existentes propostos pela proposta também são abordados. Uma análise sobre o estado da arte é feita para levantar os principais estudos feitos em IoMT assim como os resultados obtidos nestes trabalhos. São explorados ainda os conceitos de Mthings e Mwearables definidos pela proposta de padrão que são utilizados em aplicações IoMT.*

1. Introdução

Nos últimos anos, é notável o aumento do número de serviços e aplicações oferecidos pela internet. Atualmente de acordo com [Alvi 2015], há algo em torno de 9 bilhões de dispositivos conectados e um aumento explosivo deste número é esperado para a próxima década. Além disso, tem sido grande o aumento do número de dispositivos conectados diferente daqueles mais tradicionais como desktop, laptops e celulares. Outros objetos físicos agora possuem também essa capacidade de se comunicar através da internet.

Esses objetos têm capacidade de observar e interagir com o ambiente físico e de se comunicar com outros objetos sendo assim conhecidos como um novo paradigma na internet chamado de IoT (Internet of Things). IoT tem uma capacidade enorme de influenciar nossas vidas e a maneira como interagimos como objetos estando presente em muitas aplicações como automação residencial, cidades inteligentes, monitoramento e gerenciamento de dispositivos automatizados, entre muitas outras.

Em [Charu 2013], [Ma 2011] e [Atzori 2010] são feitos estudos que definem IoT em termos de capacidades de detecção e atuação de coisas, redes, tecnologias web e potenciais de mercado. Ainda englobam obstáculos para a sua padronização e avaliam o

impacto de sua utilização para a sociedade. Entretanto, estes estudos não consideram os requisitos e desafios impostos pelos dispositivos multimídia ou o transporte do tráfego multimídia pela rede com outros dados. Os dados multimídia impõem uma série de restrições temporais e espaciais que devem ser consideradas no cenário IoT. Para o envio de conteúdo multimídia de forma a não prejudicar a experiência do usuário, é necessário respeitar certos requisitos de qualidade de serviço (Quality of Service - QoS) que varia de aplicação para aplicação, mas que em geral envolve características da rede em termos de atraso fim-a-fim, jitter, taxa de erro, entre outros.

Esse aumento na utilização de dados multimídia na rede se deve principalmente à popularização de aplicações e serviços como videoconferência, vídeo sob demanda, aplicações multimídia em tempo real, jogos on-line, entre muitas outros. Isso mostra que o tráfego multimídia tende a ser o principal na Internet especialmente em termos de conteúdo de vídeo. Um estudo [Jridi 2018] mostra tendências e previsões globais sobre o tráfego na Internet e foi sinalizado um grande aumento no fluxo de tráfego multimídia nos próximos 5 anos. A Cisco [Cisco 2017] também fez uma previsão em relação a tendência do tráfego IP global na Internet de aplicações visuais na rede. A Figura 1 mostra o resultado dessa previsão e conforme os resultados apresentados, é notável uma tendência no aumento do tráfego multimídia especialmente em relação a conteúdo de vídeo nos próximos anos que deve dominar o tráfego global da Internet.

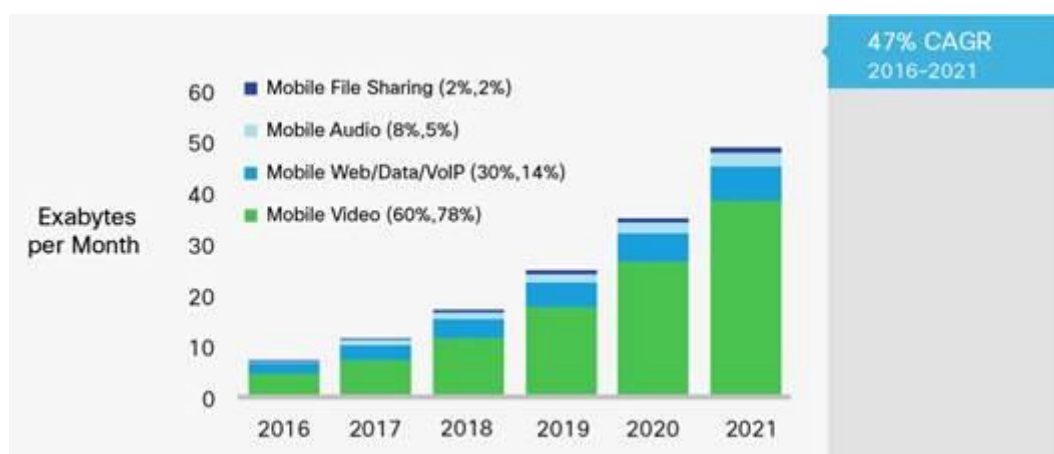


Figura 1. Uso e Previsão do Tráfego Global da Internet [Cisco 2017]

Os dados multimídia possuem características distintas daqueles escalares adquiridos por dispositivos IoT. Dispositivos multimídia requerem maiores recursos de processamento e de memória e consomem mais a largura banda em relação ao tráfego de dados convencionais nas redes IoT. Por isso, para tratar dessa funcionalidades e restrições adicionais, foi criado uma subárea em IoT chamada de (Internet of Multimedia Things - IoMT).

Desde 2014, o MPEG International Standardization Group (ISO/IEC SC29 WG11) criou um grupo chamado de (Internet of Media Things and Wearables – IoMT) com o objetivo de definir interfaces entre coisas de mídia e humanos. Esta proposta de padrão suporta a interoperabilidade de formatos de dados entre coisas de mídia incluindo sensores, atuadores, analisadores de mídia e seres humanos.

Os dados multimídia em IoMT são bem maiores em relação aos dados manipulados por aplicações IoT tradicionais e especialmente para comunicação em tempo real, maiores recursos de processamento e memória são necessários. A entrega destes dados deve satisfazer a certas restrições de QoS que obrigam a utilização de uma largura de banda maior e mecanismos de comunicação eficientes.

A comunicação multimídia através de redes sem fio foi abordada em muitos estudos, com escopos específicos dos dispositivos [Akyildiz 2007], específico da aplicação [Akyildiz 2008] e específico do conteúdo [Almalkawi 2010]. Em [Alvi 2015] no entanto, foca-se em apresentar uma arquitetura para IoMT que envolve a heterogeneidade de dispositivos multimídia, problemas de largura de banda, processamento multimídia complexo, serviços em nuvem e outros problemas.

O objetivo deste trabalho é abordar a proposta de padrão MPEG IoMT através da definição do seu objetivo e da apresentação do estado da arte em relação ao que está sendo estudado e resultados obtidos em aplicações com a utilização dessa proposta de padrão. Além de mostrar também alguns conceitos e casos de uso que a própria proposta apresenta.

O resto deste trabalho está organizado como segue. Na Seção 2 são apresentados os trabalhos relacionados. Uma visão geral sobre a proposta de padrão MPEG IoMT é feita na Seção 3. Na Seção 4 são discutidos os conceitos de mthings e mwearables e a Seção 5 expõe alguns casos de uso propostos. Por fim, na Seção 6 é feita a conclusão deste trabalho.

2. Trabalhos Relacionados

Um estudo feito [Kim 2017], descreveu alguns termos e casos de uso utilizados em IoMT que foram definidos pela proposta de padrão MPEG além de mostrar resultados de um sistema experimental de assistência de navegação para pessoas cegas. Em [Yang 2017] os autores propõem um método de detectar e reconhecer gestos manuais para gerar comandos baseados em gestos para controlar o consumo de mídia em óculos inteligentes. Para isso, os autores utilizaram um método de detecção que utiliza imagem em profundidade obtida por sequências de imagens e informações de cor da pele de forma combinada. Eles representam os contornos da mão detectados com base na curva de Bézier como metadados para fornecer uma interface interoperável entre módulo de reconhecimento em uma estrutura IoMT. A curva de Bézier é descrita por [Forrest 1972] como um segmento de curva que é definido como por um polígono cujos dois vértices são os pontos finais do arco.

Uma proposta de arquitetura para atender os requisitos de IoMT é feita em [Alvi 2015] baseada nos principais desafios levantados pelos autores pelo paradigma IoMT. Um modelo de QoE em camadas é proposto em [Floris 2015] para avaliar e combinar contribuições de cada fator de influência para estimar a QoE em aplicações IoMT. Para avaliar a abordagem proposta, eles utilizam uma aplicação IoMT veicular para aulas remotas de direção. Em [Balan 2017], os autores propõem a utilização do LISP (Locator Identifier Separation Protocol) para multihoming e balanceamento de carga. Para isto, eles utilizam um demonstrador de um gateway LISP IoT móvel que é integrado à análise de vídeo baseada em nuvem.

Uma questão bastante importante para a comunidade de tecnologia de informação e comunicação é a preocupação com o aumento das emissões de CO₂ que exige comunicação verde para reduzir o consumo de energia e as emissões de carbono. No trabalho feito em [Alvi 2015] os autores citam que o grupo IETF ROLL padronizou um protocolo de roteamento IPv6 para redes de baixa potência e perdas (Routing Protocol for Low Power and Lossy Networks - RPL) para dispositivos com recursos limitados. No entanto, implementações RPL pensando nos requisitos do tráfego multimídia não foram realizadas. Por isso neste trabalho, os autores projetam uma versão RPL para IoMT que minimiza a emissão de carbono e o consumo de energia, porém com a incorporação dos requisitos de QoS específicos da aplicação. Para avaliar a proposta, é feito um estudo de simulação no Cooja simulator para Contiki-OS que oferece ganhos em termos de eficiência energética e atraso.

Em [Li 2017] os autores apresentam uma abordagem baseada em sensoriamento comprimido que combina o método de aprendizado de dicionário (Dictionary Learning - DL) e a aproximação aproximada de passagem de mensagem (Approximate Message Passing Approach - AMP). Eles aplicam um método de denoising baseado no aprendizado de dicionários na estrutura do algoritmo AMP e propõem um framework DL-AMP. Nas experimentações feitas, demonstrações comprovaram a eficácia da estrutura para dispositivos IoMT com a capacidade de reduzir o uso da largura de banda necessária para a comunicação multimídia.

Um protocolo cross-layer para IoT Multimídia (IoMT) é apresentado em [Rani 2017]. Os autores consideraram a comunicação cruzada das camadas física, enlace e roteamento para aplicações multimídia. Eles utilizaram as camadas tradicionais do MATLAB para comparar a metodologia proposta e obtiveram um ganho no consumo energético e em minimização do atraso. Um protocolo (Power Save Multiple Poll - PSMP) do tipo (Multi-Hop Power Save Multiple Poll - mPSMP) com reconhecimento de QoS para permitir a comunicação multimídia com eficiência energética através de IoT também é proposto em [Afzal 2016]. O mPSMP possui um modelo de tráfego para alocar recursos de canal utilizando acesso múltiplo por divisão de tempo. O ciclo de serviço adaptativo é empregado para minimizar a utilização de energia enquanto assegura QoS multimídia necessária a cada nó. A análise do protocolo proposto foi implementada no Network Simulator-2 (NS-2).

Os autores do trabalho em [Saha 2018] propuseram um algoritmo de estimativa de movimento baseado em padrão adaptativo sensível ao contexto para a plataforma de IoT multimídia para melhorar a compactação de vídeo. Os movimentos no algoritmo são classificados em grandes ou pequenos com base no valor de distorção. O padrão de pesquisa é escolhido entre o padrão de pesquisa de pequenos losangos (Small Diamond Search Pattern - SDSP) ou o padrão de pesquisa de grandes losangos (Large Diamond Search Pattern - LDSP) em cada passo de estimativa de movimento permitindo assim, o processamento adaptativo de informações abstratas grandes e pequenas.

Em [Jridi 2018], é apresentado um cooprojeto de algoritmos/arquitetura e hardware/software para a implementação de uma camada de computação digital de borda em uma plataforma Zynq no contexto da Internet das Coisas Multimídia (IoMT). O sistema de computação de borda proposto é composto de um módulo reconfigurável

para comprimir e criptografar simultaneamente múltiplas imagens juntamente com a transmissão sem fio de imagens e as funcionalidades de exibição.

3. Visão Geral MPEG IoMT

Internet of Multimedia Things (IoMT) é definida como um conjunto de interfaces, protocolos e representações de informações relacionadas às mídias que permitem serviços e aplicações baseados na interação entre humano e dispositivo e dispositivo e dispositivo tanto em ambientes físicos quanto em virtuais.

Dois sistemas principais são utilizados para implantar serviços em aplicações multimídia IoT: Wireless Multimedia Systems (WMS) e Wireless Multimedia Sensor Networks (WMSN). Os sistemas WMS, são usados para monitorar o ambiente e onde os sensores podem receber feedback para controlar o processo de aquisição enquanto que os sistemas WMSN os dispositivos multimídia não têm (ou possuem pouca) capacidade para receber feedback.

Na Figura 2, é ilustrado a arquitetura típica dos sistemas WMS, onde as fontes multimídia enviam os dados na Internet através de um gateway onde a informação multimídia é processada e armazenada para acesso síncrono ou assíncrono dos usuários e dos administradores.

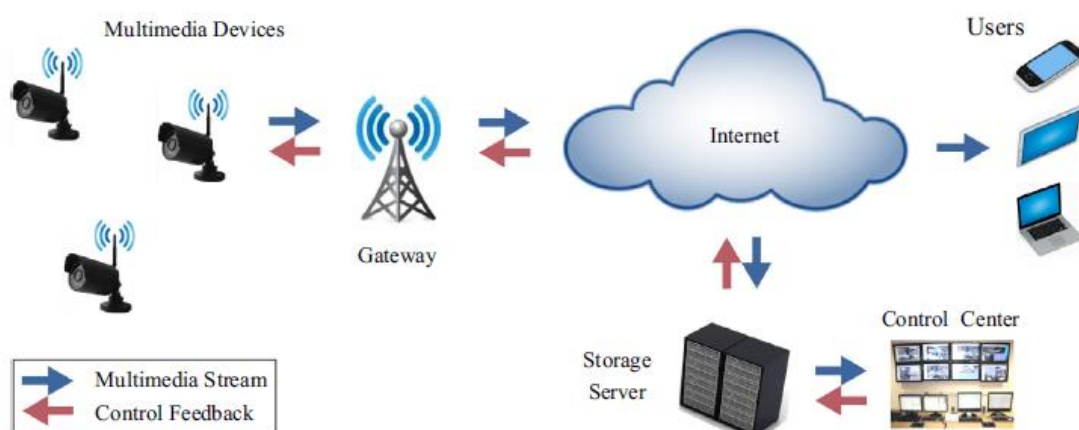


Figura 2. Arquitetura Típica de WMS [Alvi 2015]

Entretanto, apesar do sucesso dos sistemas WMS, eles possuem algumas limitações que impõem restrições quanto à sua adaptação. O escopo desses sistemas é limitado ao cenário de implantação com uma arquitetura fixa com mobilidade reduzida, um conjunto pré-definido de dispositivos e funcionalidades. Não há restrição para que as soluções tenham uma boa eficiência energética e dispositivos multimídia que possuem pilhas de comunicação semelhantes não conseguem se comunicar com outros dispositivos na rede realizando diferentes tarefas.

Na Figura 3 é mostrada a arquitetura padrão utilizada nos sistemas WMSN. As fontes multimídia em WMSN tem funcionalidades limitadas e enviam o conteúdo por meio de um WMS. Nestes sistemas, o escopo também é limitado ao cenário de implantação onde as características dos dispositivos da rede são conhecidas no momento

da implantação. Dessa forma a operação e os requisitos de QoS tanto nos dispositivos quanto na rede são conhecidos e pré-determinados. Pelos dispositivos nos sistemas WMSNs possuem pouca heterogeneidade em termos de recursos e capacidades, as operações na rede e os requisitos de QoS não conseguem ser adaptáveis aos cenários de rede atuais.

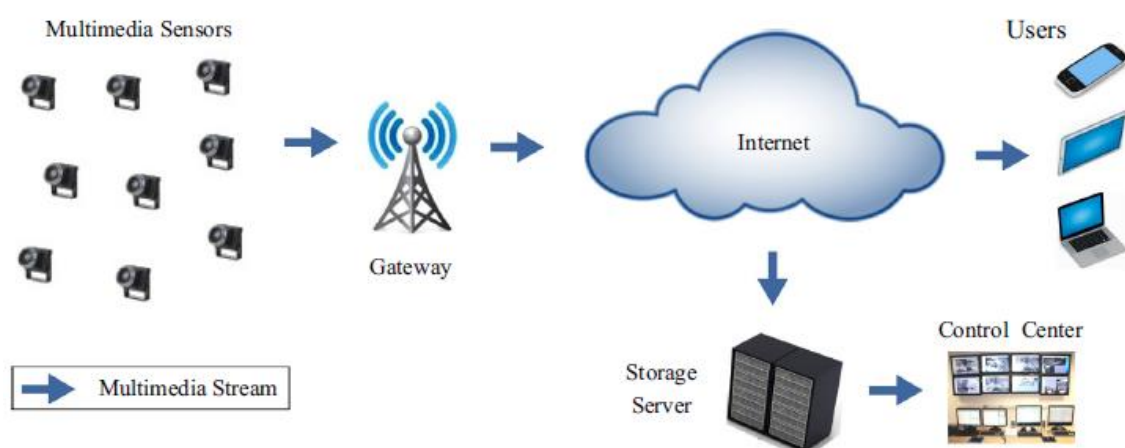


Figura 3. Arquitetura Típica de WMSN [Alvi 2015]

As aplicações IoMT possuem requisitos de QoS semelhantes às redes multimídia tradicionais para garantir uma experiência do usuário satisfatória. Algumas aplicações podem ser tolerantes a perdas enquanto outras são sensíveis ao atraso. As aplicações multimídia segundo [Kurose 2010] podem ser categorizadas em três classes: (i) transmissão de mídia contínua armazenada, (ii) transmissão de mídia contínua ao vivo e (iii) transmissão de mídia contínua interativa em tempo real.

Atualmente, para realizar as garantias de QoS exigidas pelas aplicações são utilizadas técnicas em nível de aplicação uma vez que a Internet TCP/IP prove o serviço de melhor esforço independente do conteúdo sendo transmitido não sendo capaz dessa forma, de garantir os níveis de QoS necessários para cada aplicação. Além disso, diferentemente dos sistemas IoT onde a largura de banda utilizada é geralmente baixa, nos sistemas IoMT a quantidade de dados gerados é variável podendo exigir que a largura de banda oferecida seja adaptável.

A proposta de padrão MPEG IoMT prevê que dispositivos multimídia como câmeras e microfones sejam acessíveis globalmente através de um endereço IP exclusivo assim como outros dispositivos de rede conectados à Internet. Devido a este conceito, dispositivos multimídia heterogêneos que obtêm informações multimídia do ambiente físico podem se comunicar e interagir uns com os outros, assim como outras coisas (“things”) inteligentes conectadas através da rede global (Internet).

Em [Alvi 2015], IoMT é definido como uma rede global de coisas multimídia que são exclusivamente identificados e endereçados para adquirir dados multimídia, bem como possuem capacidades de interagir e se comunicar com outros dispositivos e serviços multimídia ou não, com ou sem intervenção humana. Ainda é proposta uma

arquitetura para IoMT de quatro estágios para descrever o funcionamento de um serviço baseado em IoMT para solucionar as limitações dos sistemas WMS e WMSN.

A Figura 4 ilustra a arquitetura proposta onde o primeiro estágio consiste em adquirir e codificar o conteúdo multimídia, no segundo, o conteúdo é relatado a nuvem onde são incorporadas técnicas de comunicação e endereçamento. Após isso no terceiro estágio, a informação multimídia é armazenada, processada e disseminada de acordo com a demanda dos usuários finais e por fim, no quarto estágio, são executadas as tarefas computacionais de pós-processamento na nuvem de acordo com os requisitos de aplicação/serviço.

Um estudo feito em [Floris 2015], mostrou que em IoMT podem existir três cenários distinguidos com base no uso do conteúdo multimídia. Um deles é a multimídia como uma entrada IoT onde o conteúdo multimídia é adquirido e os objetos multimídia são usados por uma aplicação IoT para fornecer um determinado serviço. Outro cenário diz respeito à multimídia como saída IoT onde objetos IoT adquirem sinais, dados e informações (não multimídia) e apresentam-nos de forma multimídia através de uma aplicação IoT. Por fim, ainda existe um cenário onde a multimídia é a entrada e a saída de IoT, o conteúdo é adquirido por objetos multimídia e é apresentado de forma multimídia por uma aplicação IoT.

Na Figura 5, é ilustrado um exemplo de aplicação para cada um dos cenários acima. O primeiro cenário (representado pela cor vermelha) mostra uma aplicação onde câmeras são utilizadas para registrar imagens de pessoas que desejam entrar em um local onde a entrada é permitida somente a pessoas autorizadas. As imagens gravadas são enviadas para uma plataforma IoT onde é implantada uma aplicação que consiste de um software de identificação para determinar se a pessoa gravada na câmera está autorizada a acessar o local. Assim que o software identificar a pessoa um comando é enviado para o atuador da porta abri-la caso a pessoa esteja autorizada. Caso contrário, a porta permanece fechada e o acesso é negado.

O segundo cenário (representado pela cor verde) ilustra uma aplicação onde vários objetos IoT medem alguns parâmetros médicos de um paciente (temperatura, pulso, pressão, etc) que são coletados por uma plataforma IoT. A aplicação tem a tarefa de apresentar o status do paciente de uma maneira multimídia através de gráficos, animações, alarmes, etc. No terceiro cenário (representado pela cor laranja), algumas câmeras de vigilância gravam imagens e áudio de um lugar. A informação multimídia é coletada por uma aplicação IoT e apresentada de forma multimídia para fornecer um serviço de controle remoto de segurança.

Em [ISO/IEC TC1/SC29/WG11 N17679 2018] é determinado através de uma especificação do grupo MPEG IoMT que as aplicações IoMT devem atender a certos requisitos de detecção e aquisição de dados, processamento de dados, tempo, atuação, comunicação, renderização e apresentação, armazenamento, segurança e privacidade, agregação do sistema e de token de mídia que são tokens gerados ao utilizar aplicações multimídia com blockchains. As aplicações que utilizam blockchains devem especificar a lista de blockchains suportados. Todos os requisitos com suas respectivas descrições podem ser encontrados em [ISO/IEC TC1/SC29/WG11 N17679 2018].

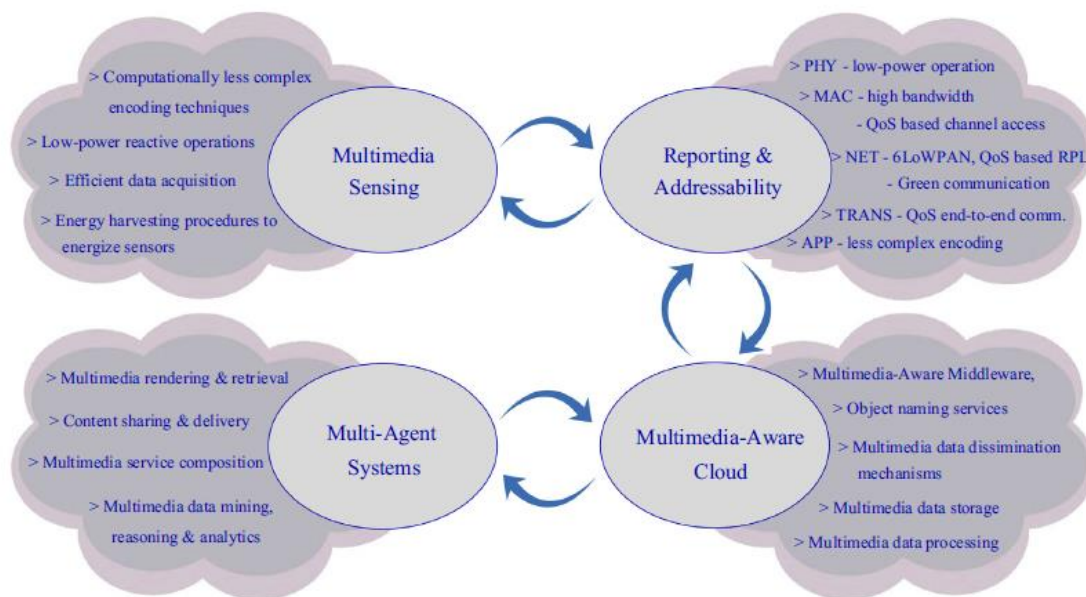


Figura 4. Arquitetura IoMT proposta [Alvi 2015]

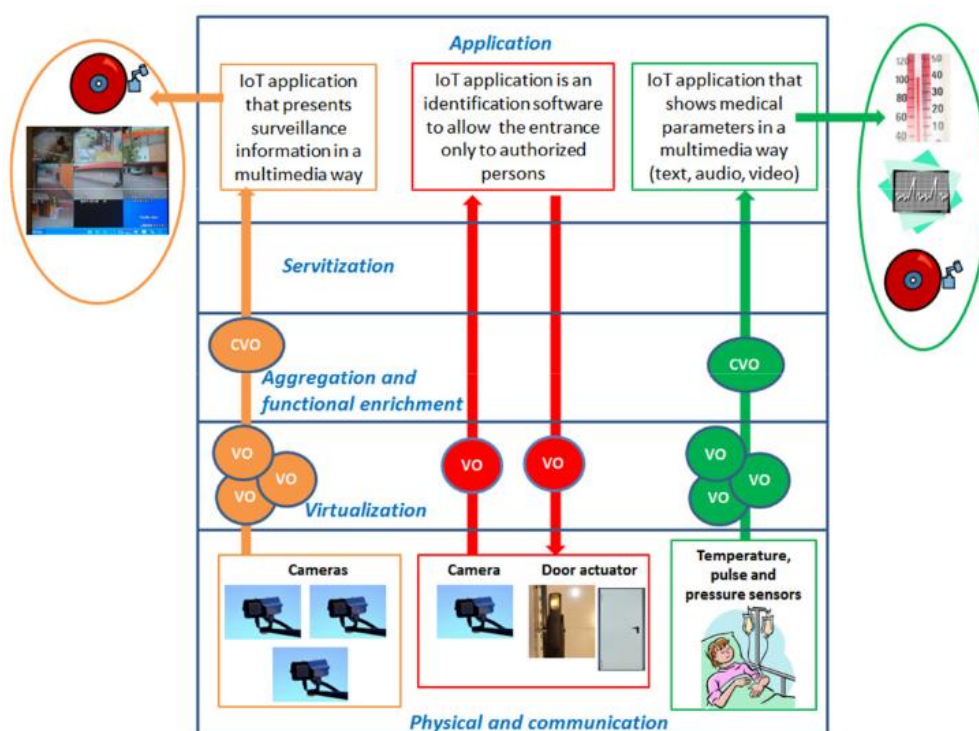


Figura 5. Cenários IoMT [Floris 2015]

4. Mthings

Para definir o que é uma media thing (mthing) em [Kim 2015], a proposta de padrão MPEG IoMT primeiramente especifica alguns conceitos como o de entidade que é

qualquer objeto físico ou virtual que é percebido e/ou influenciado pelas things. Uma thing é qualquer coisa que pode se comunicar com outras além de poder sentir e/ou agir em entidades. Por fim, media thing é uma coisa com pelo menos um dos recursos de detecção e ativação áudio/visual. Na Figura 6 é mostrado um esquema que representa a forma como os componentes descritos acima estão relacionados.

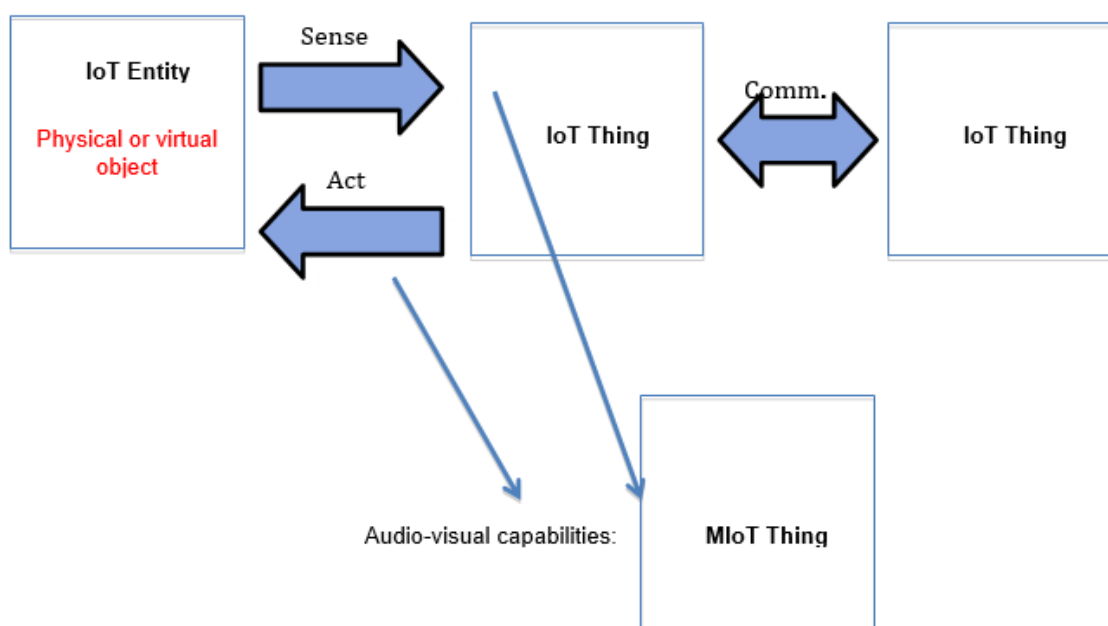


Figura 6. Representação Mthings [Kim 2015]

MPEG IomT ainda especifica que as media things podem executar as tarefas de detecção de movimento e classificação de objetos, cooperação com outras visões das coisas para rastreamento de objetos e compressão e transmissão de áudio/vídeo. Elas devem ter o desempenho de monitorar sua área de observação com rapidez suficiente para detectar e analisar todos os objetos de interesse em tempo real, codificar áudio/vídeo com eficiência energética e expor parâmetros de QoS e QoE. Para atender os requisitos relacionados às funcionalidades audiovisuais do IoMT, algumas ferramentas MPEG são utilizadas conforme mostra a Tabela 1.

Tabela 1. Ferramentas MPEG relacionadas a IoMT [Kim 2015]

Requirements	MPEG Tool
Motion detection	MPEG-7
Object detection and classification	CDVS (and probably CDVA), some MPEG-7 descriptors
Cooperation with other vision Things for the object tracking.	MPEG-4 Systems
Synchronization of data from all sensors	MPEG-4 Systems, MPEG-V
Persistence of metadata among processing chain	MPEG-7 (probably extended)
Energy efficiency for audio / video compression and transmission	MPEG-4 Video, Audio

Alguns exemplos de Mthings como câmera digital, microfone inteligente, unidade de processamento de mídia e óculos inteligentes são descritos em [Kim 2017]. A câmera digital é um dispositivo IoMT cuja funcionalidade principal é a de uma câmera. Além disso, pode ter também funcionalidades adicionais como compressão, armazenamento e transmissão ou streaming de dados visuais adquiridos.

Microfone inteligente é um dispositivo IoMT cuja funcionalidade principal é a de um microfone. Pode haver funcionalidades adicionais como compressão, armazenamento e transmissão ou streaming de dados acústicos adquiridos. Unidade de processamento de mídia é um dispositivo IoMT cuja funcionalidade principal é o processamento de mídia ou metadados detectados ou adquiridos

A Unidade de Processamento de Mídia pode ter funcionalidades adicionais como reconhecer e interpretar gestos, reconhecer e/ou interpretar sons e extrair contornos e regiões. Além disso, é capaz de fornecer descrições áudio/vídeo espaciais e temporais e reconhecer e interpretar descritores semânticos relacionados a determinados eventos (por exemplo, incêndio, colisão, luta, intrusão, etc.). Esta unidade fornece metadados relacionados a eventos para detectar objetos de interesse (por exemplo, pessoa, carro, logotipos) e fornece descritores a eles para comparar os descritores dos objetos detectados com os descritores no armazenamento local para gerar metadados semânticos dos objetos detectados (por exemplo, sexos, idades, tipos de carro, matrículas). Este dispositivo ainda é utilizado para sintetizar sons, para traduzir idiomas e gerar respostas de linguagem natural.

Há ainda um subconjunto de dispositivos em IoMT descritos pela proposta de padrão MPEG em [Mitre 2016] que são os chamados wearable devices que são aqueles que podem ser usados por ou incorporados em uma pessoa e têm a capacidade de se conectar e comunicar diretamente com a rede através de conectividade celular ou através de outro dispositivo usando Wi-Fi, Bluetooth ou outra tecnologia. Alguns exemplos desses dispositivos que normalmente são usados para se comunicar com os serviços em nuvem por meio de smartphone e smart TVs são mostrados na Figura 7.

Os óculos inteligentes conforme descritos em [Yang 2017], são MWearables cuja funcionalidade principal é fornecer uma interface interoperável entre o usuário e o mundo físico/virtual. Os óculos inteligentes possuem uma gama de funcionalidades, incluindo aquelas associadas a determinados dispositivos como câmera, microfone, processamento de mídia, exibição e alto-falante. Além disso, óculos inteligentes devem ser capazes de fornecer a Interface Natural do Usuário (NUI) (por exemplo, gesto com a mão, movimento da cabeça, gesto corporal, voz, marcador, rastreamento ocular) e fornecer NUI de forma combinada (multimodal) e de forma individual (por exemplo, gestos ou vozes).



Figura 7. Exemplos Wearable Devices [Mitrea 2015]

5. Casos de Uso IoMT

Os casos de uso mostrados nesta seção foram explorados pelo MPEG IoMT em [Mitrea 2016] e são categorizados em espaços inteligentes de monitoramento, navegação e ambiente inteligente. Além dos discutidos aqui, ainda existem inúmeros casos de uso uma vez que qualquer combinação de itens da mídia é viável para oferecer suporte a um novo tipo de serviços de mídia IoT.

5.1 Espaços Inteligentes: Monitoramento e controle com rede de câmeras de áudio-vídeo

5.1.1 Reconhecimento facial para invocar atuações sensoriais

Nesta aplicação, uma câmera de vigilância IP (câmera IoMT) captura dados de áudio/vídeo e os envia para um armazenamento (armazenamento IoMT) e uma unidade reconhecadora de face (unidade de processamento de mídia). Quando o reconhecedor de rosto detecta e reconhece a face de uma pessoa pré registrada, ele ativa um gerador de aroma para borrifar um perfume específico. Os descritores específicos (como locais de rosto detectados, descritores de face, locais de mídia de momentos detectados) podem ser extraídos e enviados para um armazenamento conforme ilustrado na Figura 8. Neste caso de uso, o gerador de perfume pode ser substituído por qualquer tipo de atuador (como lâmpadas, displays, tocadores de música, etc).

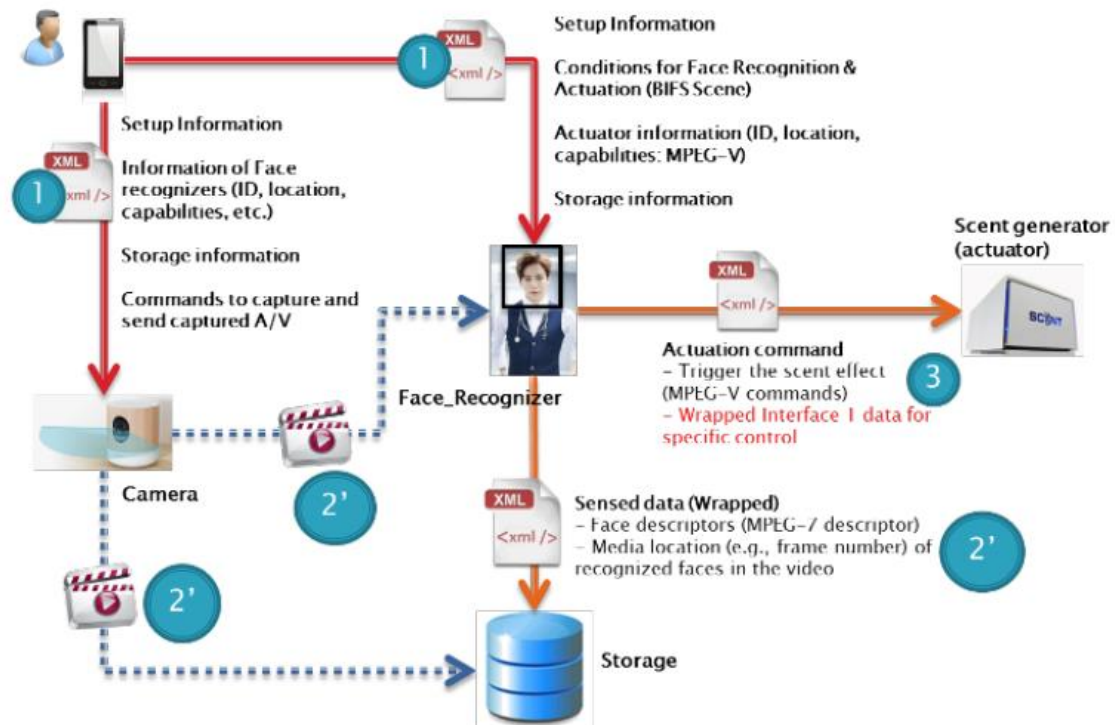


Figura 8. Esquema de fluxo de dados invocando atuações sensoriais usando detecção de faces [Mitrea 2016]

5.1.2 Rastreamento humano com várias câmeras de rede

Esta aplicação descreve um sistema de vídeo-vigilância utilizado em cidades para reduzir o número de crimes urbanos. O sistema inclui um conjunto de câmeras de vigilância IP (câmeras IoMT), um armazenamento (storage IoMT) e uma unidade de rastreamento humano (unidade de processamento de mídia) como ilustrado na Figura 9.

Uma determinada câmera de vigilância IP captura os dados de áudio/vídeo e os envia para o armazenamento e para a unidade rastreadora humana. Quando o rastreador humano detecta uma pessoa na área visível, rastreia a pessoa e extrai a trajetória em movimento. Caso a pessoa fique fora do alcance visual da primeira câmera IP, mas permaneça na área protegida pelo sistema de vídeo-vigilância da cidade, outra câmera IP desse sistema nas proximidades assumirá o controle e continuará capturando dados de áudio/vídeo da pessoa.

Se a pessoa sair da área protegida pelo sistema de vídeo-vigilância da cidade, por exemplo, a pessoa entra em um centro comercial, então o sistema de busca da cidade checka se este centro comercial também está equipado com um sistema de vigilância por vídeo. Se estiver, o sistema de vídeo-vigilância da cidade estabelece uma comunicação com o sistema de vídeo vigilância do centro comercial para permitir que outra câmera IP do centro de vigilância de vídeo comercial continue a capturar dados de áudio/vídeo da pessoa correspondente. Neste caso, as interfaces entre o sistema de vídeo-vigilância da cidade e o sistema de vídeo-vigilância do centro comercial devem ser padronizadas para comunicar e trocar dados interoperáveis. Em ambos os casos, os descritores específicos (como informações de trajetória em movimento, aparência das informações, locais de mídia de momentos detectados) podem ser extraídos e enviados

para o armazenamento. Neste caso de uso, o módulo rastreador humano pode controlar a ativação e desativação de câmeras na área.

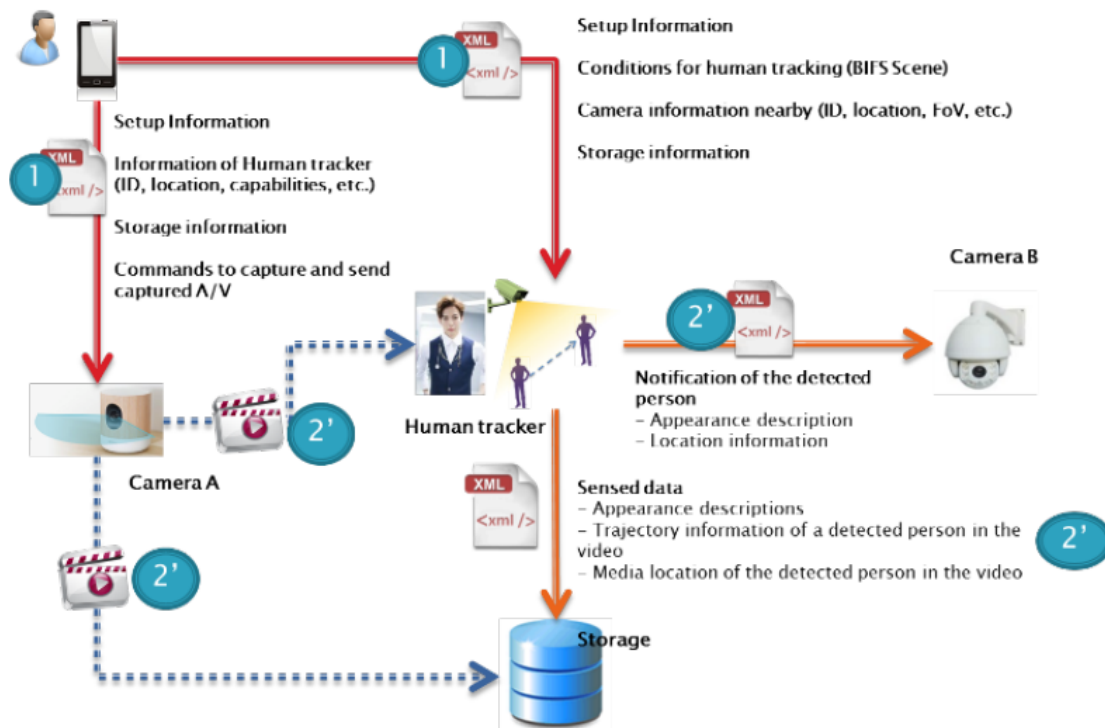


Figura 9. Esquema de fluxo de dados de um sistema de vídeo vigilância de uma cidade [Mitrea 2016]

5.2 Espaços Inteligentes: Navegação

5.2.1 Aviso de colisão

Esta aplicação tem o objetivo de evitar que um pedestre cego colida com um obstáculo a sua frente através de dispositivos IoMT conforme mostra a Figura 10. Neste cenário a pessoa carrega uma bengala inteligente, uma faixa vibratória, um smartphone e um fone de ouvido em rede. A bengala inteligente equipada com sensores de distância (por exemplo, um sensor ultrassônico ou infravermelho) pode medir a distância entre a bengala e os obstáculos a frente. Um coordenador de colisão (unidade de processamento de mídia) recebe estes dados e decide quais ações a serem executadas. Se a distância for razoavelmente longa, um texto de alarme de distância correspondente é produzido pelo coordenador de colisão e enviado para uma unidade geradora de conversão de texto em voz (unidade de processamento de mídia). O gerador cria o arquivo de áudio correspondente e envia sua URL para um fone de ouvido em rede. O fone reproduz os arquivos de áudio para a pessoa cega possibilitando-a desviar do obstáculo. Se a distância estiver muito perto, o coordenador de colisão ativa a pulseira vibratória ou o fone de ouvido para criar sons de bipe. O esquema de fluxo deste caso de uso é ilustrado na Figura 11.



Figura 10. Sistema de assistência a pessoas cegas para evitar obstáculos [Mitrea 2016]

5.2.2 Direção de orientação

A Figura 12 mostra o caso de uso onde uma pessoa cega deseja chegar a um destino. Para isto, são utilizadas etiquetas RFID (ou beacons) que possuem as coordenadas exatas da localização. Esta pode ser registrada e recuperada pelo servidor usando seus IDs exclusivos. As etiquetas dessa forma podem estar em todos os cantos da rua. A pessoa cega carrega uma bengala inteligente, um smartphone e um fone de ouvido em rede. A bengala possui um leitor de RFID e alguns sensores de inércia (como um giroscópio ou uma bússola). O leitor de RFID pode ler as etiquetas RFID em todas as esquinas das ruas. Um direcionador de movimento (unidade de processamento de mídia) recebe os dados da etiqueta RFID (ou dados do beacon) e recupera a localização atual da pessoa.

Através da combinação com outras informações de inércia, o direcionador cria uma orientação direcional (por exemplo, “vire à esquerda” ou “siga em frente”) e envia para uma unidade geradora de texto para fala (unidade de processamento de mídia). Após fazer a conversão e criar o arquivo de áudio correspondente, o gerador envia a URL correspondente para um fone de ouvido em rede. O fone assim reproduz os arquivos de áudio e indica a direção que a pessoa cega deve seguir conforme mostra a Figura 13.

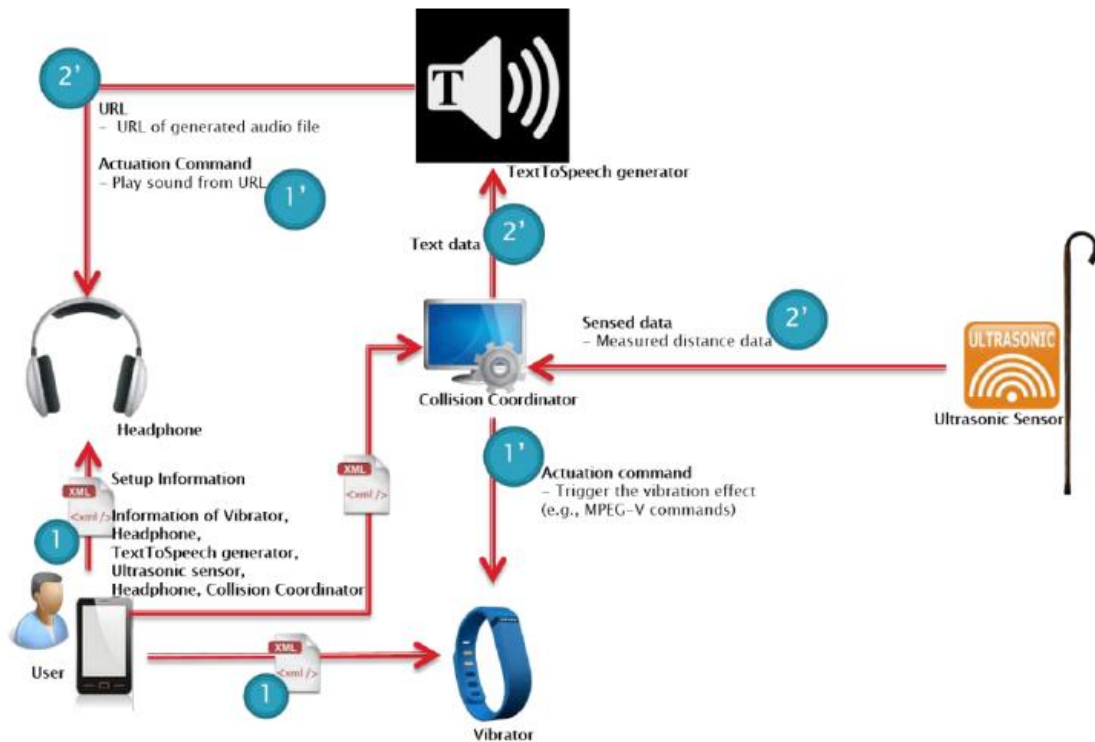


Figura 11. Esquema de fluxo de dados de um sistema de prevenção de colisão para pessoas cegas [Mitrea 2016]

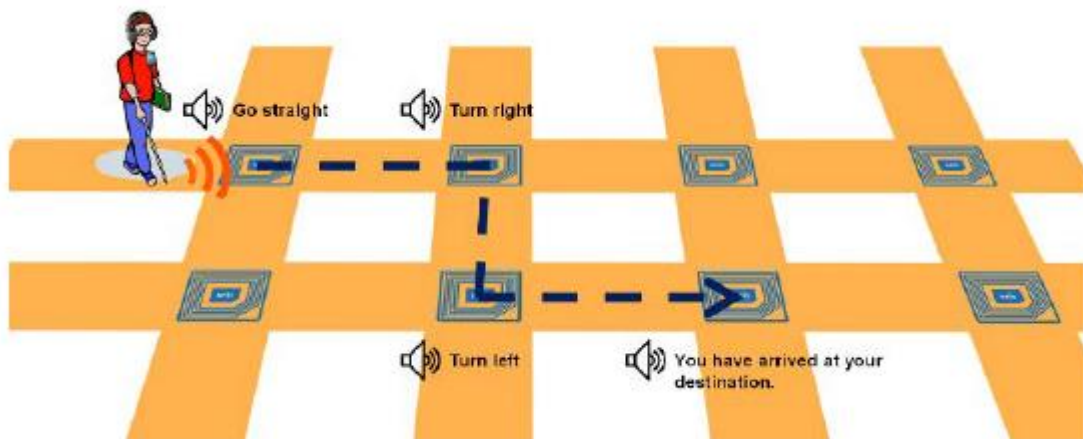


Figura 12. Cenário de uso de um sistema de navegação para pedestres cegos [Mitrea 2016]

5.2.3 Navegação turística personalizada com funcionalidades de linguagem natural

Esta aplicação consiste em um serviço de tradução de fala quando duas pessoas falam idiomas diferentes. Quando um usuário fala a um microfone em um relógio inteligente em um idioma com outro usuário com idioma diferente, sua fala será traduzida para o

idioma de destino. O resultado da tradução pode assim ser ouvido pelo usuário do idioma de destino através de um dispositivo wearable. O mecanismo de tradução fica em servidor remoto (sistema de tradução remota) ou no smartphone (sistema de tradução independente) que é conectado ao dispositivo wearable. Com este serviço, o usuário pode usar suas mãos livremente enquanto a conversa é traduzida. Além disso, o dispositivo wearable também pode ser utilizado para encontrar alguém que possa falar automaticamente um dos idiomas que o sistema de tradução embutido manipula em uma situação de viagem. O esquema de fluxo de dados deste sistema é ilustrado na Figura 14.

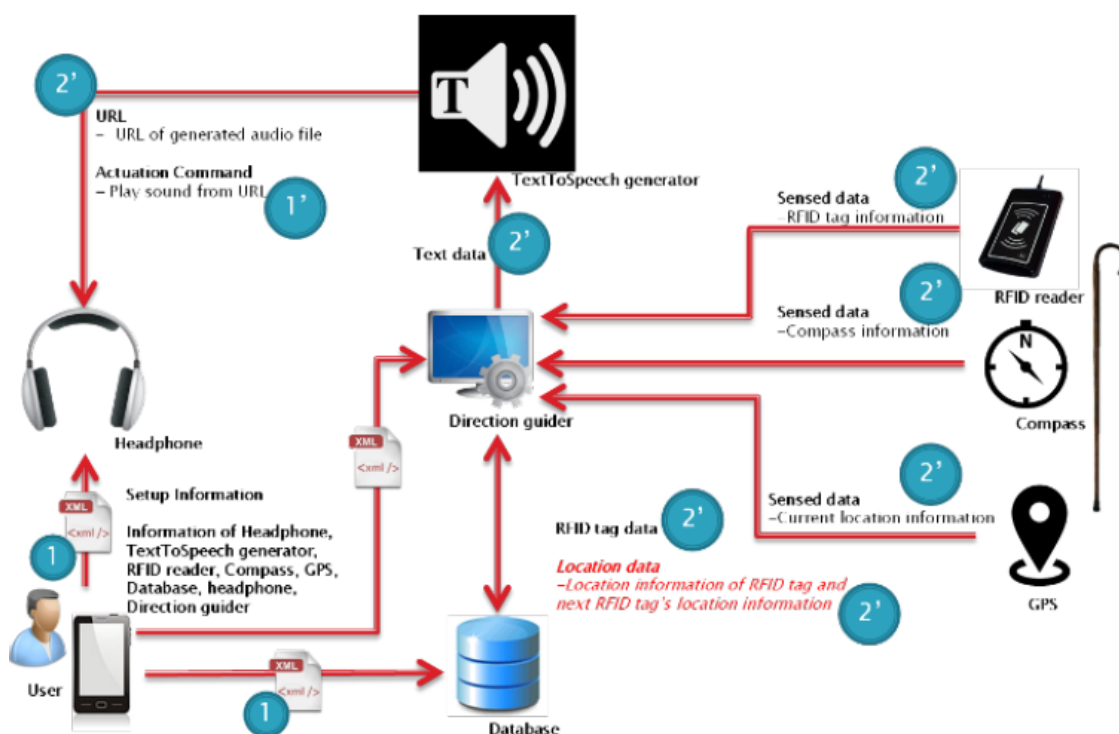


Figura 13. Esquema de fluxo de dados de um sistema de navegação para pedestres cegos [Mitreá 2016]

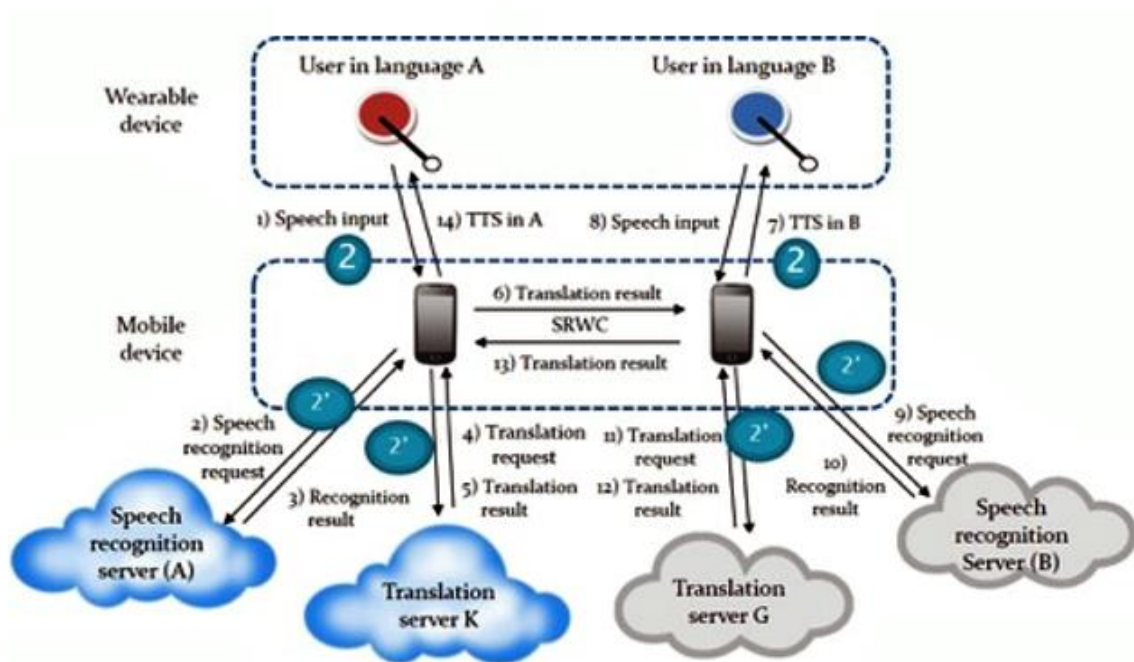


Figura 14. Esquema de fluxo de dados de um sistema de navegação turística personalizado [Mitrea 2016]

5.3 Ambientes Inteligentes em Cidades Inteligentes

5.3.1 Fábrica inteligente: sistema assistente de manutenção de carro usando óculos inteligentes

A Figura 15 ilustra o funcionamento desta aplicação onde um técnico utilizando um óculos inteligente está trabalhando na manutenção de um carro. Os óculos fornecem automaticamente uma lista de manuais de manutenção relacionados a peça específica a ser verificada no monitor. Em seguida, o técnico seleciona e lê o manual usando gestos com as mãos. Desta forma, o técnico pode realizar o trabalho de manutenção livremente utilizando as duas mãos.

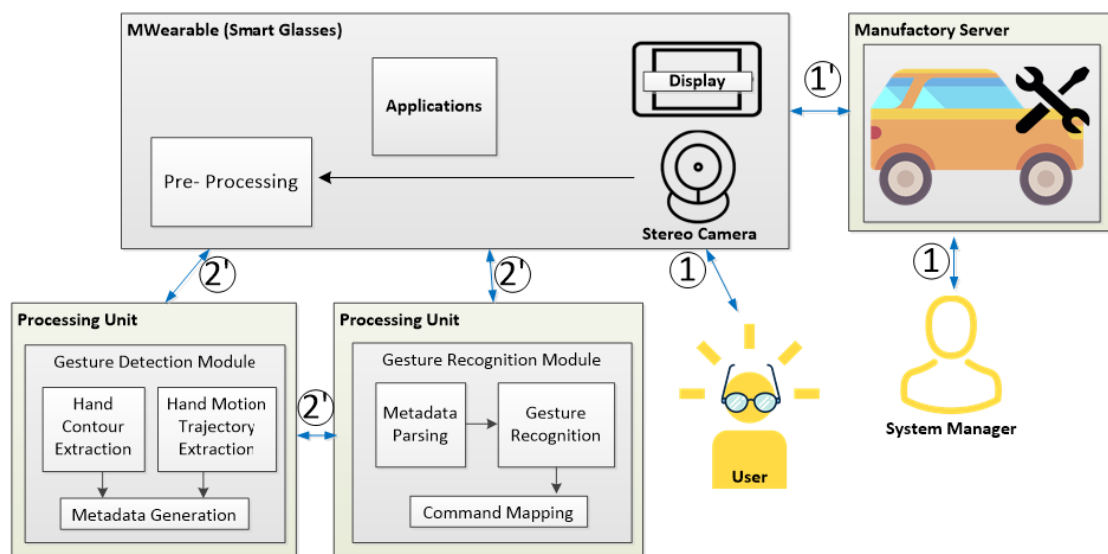


Figura 15. Esquema de fluxo de dados de um sistema assistente de manutenção de carro [Mitrea 2016]

5.3.2 Museu inteligente: educação aumentada utilizando óculos inteligentes

A Figura 16 ilustra o funcionamento deste caso de uso onde o usuário pode solicitar, através de um gesto manual, informações adicionais sobre uma determinada obra como uma explicação narrativa sobre a obra em questão ou um videoclipe mostrando uma entrevista com o pintor. Desta forma, o usuário consegue desfrutar do passeio pelo museu com informações ricas apresentadas por óculos inteligentes sem a necessidade de um folheto guia e/ou ajuda de guias.

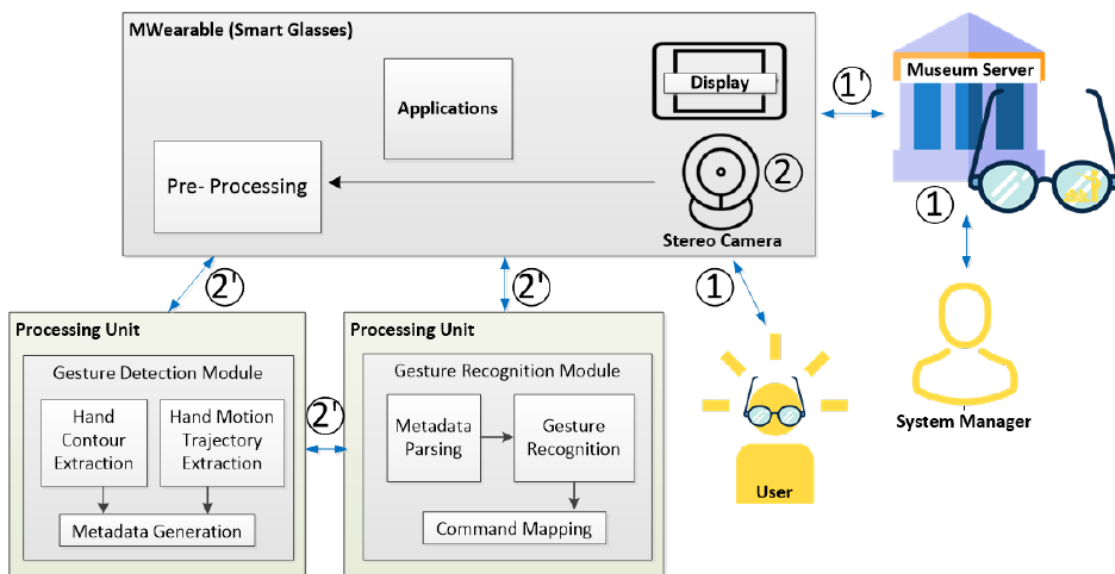


Figura 16. Esquema de fluxo de dados de um tour de realidade aumentada em um museu [Mitrea 2016]

6. Conclusão

Neste trabalho foi explorado a proposta de padrão MPEG IoMT que foi criada com o intuito de definir interfaces entre coisas de mídia e humanos. Ele especifica que as coisas de mídia (sensores, atuadores, entre outras) devem suportar interoperabilidade de formatos de dados e prevê que os dispositivos multimídia devem ser acessados globalmente através de um endereço IP exclusivo assim como outros dispositivos de rede conectados à Internet.

Devido a este princípio, dispositivos multimídia heterogêneos podem se comunicar e interagir uns com os outros assim como outras coisas (“things”) inteligentes conectadas através da rede global (Internet).

Conforme abordado na seção de trabalho relacionados, há uma série de estudos feitos recentemente que mostram a preocupação com vários aspectos de IoMT desde propostas a novas arquiteturas a protocolos que garantem uma maior eficiência energética juntamente com as garantias de QoS específicas para cada aplicação.

Além disso, foram apresentados dois sistemas principais mais utilizados para implantar serviços em aplicações multimídia IoT que são os Wireless Multimedia Systems (WMS) e Wireless Multimedia Sensor Networks (WMSN). O conceito de mthings e mwearables utilizados em aplicações IoT multimídia assim como alguns casos de uso definidos pelo MPEG IoMT utilizando estes dispositivos também foram abordados neste trabalho.

Referências

Afzal B., Alvi. S., Shah G. (2016) “Adaptative Duty Cycling based Multi-hop PSMP for Internet of Multimedia Things”, IEEE Annual Consumer Communications & Networking Conference. Disponível em: https://www.researchgate.net/profile/Bilal_Afzal4/publication/297324609_Adaptive_Duty_Cycling_based_Multi-hop_PSMP_for_Internet_of_Multimedia_Things/links/5a0d3529a6fdcc39e9bfc975/Adaptive-Duty-Cycling-based-Multi-hop-PSMP-for-Internet-of-Multimedia-Things.pdf

Alvi S., Afzal B., Shah G., Atzori L., Mahmood W. (2015) “Internet of Multimedia Things Vision and Challenge”, Ad Hoc Networks. Disponível em: https://www.researchgate.net/profile/Sheeraz_Alvi/publication/276075985_Internet_of_Multimedia_Things_Vision_and_Challenges/links/5ac7765baca272abdc5ce387/Internet-of-Multimedia-Things-Vision-and-Challenges.pdf

Alvi S., Shah G., Mahmood W. (2015) “Energy Efficient Green Routing Protocol for Internet of Multimedia Things”, IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing. Disponível em: https://www.researchgate.net/profile/Ghalib_Shah/publication/281773681_Energy_Efficient_Green_Routing_Protocol_for_Internet_of_Multimedia_Things/links/55f7c13908ae07629dcb6f12/Energy-Efficient-Green-Routing-Protocol-for-Internet-of-Multimedia-Things.pdf

Balan T., Robu D., Sandu F. (2017) “Multihoming for mobile Internet of multimedia things”, Mobile Information Systems. Disponível em: [downloads.hindawi.com/journals/misy/2017/6965028.pdf](https://www.hindawi.com/journals/misy/2017/6965028.pdf)

Charu C. Aggarwal, Naveen Ashish, Amit Sheth (2013) “The internet of things: a survey from the data-centric perspective”, in: Managing and Mining Sensor Data, Springer, pag. 383–428. Disponível em: https://www.researchgate.net/profile/Amit_Sheth/publication/235343161_The_Internet_of_Things_A_Survey_from_the_Data-Centric_Perspective/links/02e7e530dff49d62e7000000/The-Internet-of-Things-A-Survey-from-the-Data-Centric-Perspective.pdf

Cisco (2017) “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper”. Disponível em: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>

Floris A., Atzori L. (2015) “Quality of Experience in the Multimedia Internet of Things: definition and practical use cases”. Communication Workshop. Disponível em: https://www.researchgate.net/profile/Alessandro_Floris/publication/278675919_Quality

[_of_Experience_in_the_Multimedia_Internet_of_Things_Definition_and_practical_use-cases/links/5582d06308ae6cf036c2f818.pdf](#)

Forrest A. (1972) “Interactive interpolation and approximation by Bézier polynomials”, *Computer Journal*. Disponível em: <https://academic.oup.com/comjnl/article-pdf/15/1/71/1160727/15-1-71.pdf>

Hua-Dong Ma (2011) “Internet of things: objectives and scientific challenges”, *J. Comput. Sci. Technol.* 26 (6) pag. 919–924. Disponível em: <http://vote.cast.org.cn/n35081/n11219166/n12620949/n12621141/n13616780.files/n13617157.pdf>

Ian F. Akyildiz, Tommaso Melodia, Kaushik R. Chowdhury (2007) “A survey on wireless multimedia sensor networks”, *Comput. Netw.* 51 (4) pag. 921–960. Disponível em: <http://home.deib.polimi.it/cesana/teaching/IoT/2017/papers/intro/akyildiz06.pdf>

Ian F. Akyildiz, Tommaso Melodia, Kaushik R. Chowdhury (2008) “Wireless multimedia sensor networks: applications and testbeds”, *Proc. IEEE* 96 (10) pag. 1588–1605. Disponível em: http://www.eng.buffalo.edu/wnesl/people/scott/Papers/WMSN_testbed.pdf

Islam T. Almalkawi, Manel Guerrero Zapata, Jamal N. Al-Karaki, Julian Morillo-Pozo (2010) “Wireless multimedia sensor networks: current trends and future directions, *Sensors*”, 10 (7) pag. 6662–6717. Disponível em: <https://www.mdpi.com/1424-8220/10/7/6662/html>

ISO/IEC TC1/SC29/WG11 N17679 (2018) “Requirements for Internet of Media Things”, San Diego. Disponível em: <https://mpeg.chiariglione.org/standards/mpeg-iomt/iomt-architecture/requirements-internet-media-things>

Jridi M., Chapel T., Dorez V., Bougeant G., Botlan A. (2018) “SoC-Based Edge Computing Gateway in the Context of the Internet of Multimedia Things: Experimental Platform”, *Journal of Low Power Electronics and Applications*. Disponível em: <https://www.mdpi.com/2079-9268/8/1/1/html>

Kim S., Preda M., Piran Md. (2015) “Exploration on Media-centric Internet of Things (draft)”, ISO/IEC JTC1/SC29/WG11 N15082, Geneva. Disponível em: <https://mpeg.chiariglione.org/standards/mpeg-iomt/iomt-architecture>

Kurose J., Ross K. (2010), *Redes de Computadores e a Internet: uma abordagem top-down*, 5ª edição

Li Z., Huang H., Misra S. (2017) “Compressed Sensing via Dictionary Learning and Approximate Message Passing for Multimedia Internet of Things”, *IEEE Internet of Things Journal*. Disponível em: http://www.projectsgoal.com/download_projects/internet-of-things/internet-of-things-projects-GIT00022.pdf

Luigi Atzori, Antonio Iera, Giacomo Morabito (2010) “The internet of things: a survey”, *Comput. Netw.* 54 (15) (2010) pag. 2787–2805. Disponível em: <https://www.cs.mun.ca/courses/cs6908/IoT-Survey-Atzori-2010.pdf>

Mitrete M., Chun S., Pau D., Choi M. (2015) "Exploration on Wearable MPEG", ISO/IEC JTC1/SC29/WG11 w15200, Geneva. Disponible em: <https://mpeg.chiariglione.org/standards/mpeg-iomt/mpeg-wearable>

Mitrete M., Kim S., Chun S. (2016) "Use cases for Internet of Media-Things and Wearables", ISO/IEC JTC1/SC29/WG11 N16345, Geneva. Disponible em: <https://mpeg.chiariglione.org/standards/exploration/internet-media-things-and-wearables/use-cases-internet-media-things-and-1>

Rani S., Ahmed S., Talwar R., Malhotra J., Song H. (2017) "IoMT: A Reliable Cross Layer Protocol for Internet of Multimedia Things", IEEE Internet of Things Journal. Disponible em: <https://fardapaper.ir/mohavaha/uploads/2017/09/IoMT-A-Reliable-Cross-Layer-Protocol-for.pdf>

Saha A., Lee Y., Hwang Y., Psannis K. (2018) "Context-aware block-based motion estimation algorithm for multimedia internet of things (IoT) platform", Personal and Ubiquitous Computing. Disponible em: https://www.researchgate.net/profile/Kostas_Psannis/publication/317840992_Context-aware_Block-based_Motion_Estimation_Algorithm_for_Multimedia_Internet_of_Things_IoT_Platform/links/5a44c9b7a6fdcce19718f35e/Context-aware-Block-based-Motion-Estimation-Algorithm-for-Multimedia-Internet-of-Things-IoT-Platform.pdf

Sandvine (2018) "Sandvine Global Internet Phenomena Complete". Disponible em: <https://www.sandvine.com/hubfs/downloads/phenomena/2018-phenomena-report.pdf>

Sang Kyun Kim, Nevadita Sahu, Marius Preda (2017) "Beginning of a new standard: Internet of Media Things", KSII Transactions on internet and information systems, páginas 5182 – 5199. Disponible em: <http://www.itiis.org/digital-library/manuscript/1836>

Yang A., Kim J., Chun S. (2017) "Detection and Recognition of Hand Gesture for Wearable Applications in IoMT", ICACT Transactions on Advanced Communications Technology, vol. 6, issue 5. Disponible em: <https://ieeexplore.ieee.org/document/8323932/>