Vol.2019-DPS-181 No.2
2019/12/26

スポーツイベントのライブストリーミングにおける多段強化学習に基づく映像品質制御手法の提案

生出 真人[1,a)]　阿部 亨[1,2,b)]　菅沼 拓夫[1,2,c)]

概要：スポーツイベントが開催されるスタジアムではICT化が進み，観客は自身の携帯端末で選手のリアルタイム映像を視聴できるようになりつつある．しかし，端末が高密度に存在するスタジアムでの通信環境において，低遅延かつ高品質なスポーツ映像を配信するためには，ネットワーク資源の状況を考慮しつつ，端末の利用可能な資源量に適した映像品質を決定する必要がある．そこで本研究では，スポーツイベントのライブストリーミングサービスにおける映像品質制御手法を提案する．具体的には，観客が持つ端末の利用可能な資源量に適応した映像品質を効率的に決定するための，多段強化学習に基づく手法を提案する．本稿では，初期段階の各学習手法とそれらの連携方式について設計する．

1. はじめに

野球やサッカー等のスポーツイベントが開催されるスタジアムにおいて，ICT技術を駆使してファンエンゲージメントを強化した「スマートスタジアム」化が進んでいる [1]．スマートスタジアムにより，スポーツ観戦の魅力を高め，今までにない体験と感動を与えることが期待されている．スタジアム内の限定コンテンツにアクセスするため，観客は，スタジアム内のWi-Fiに自身の携帯端末を接続し，専用のアプリケーションをインストールする．このアプリケーションを利用することにより，スタジアムにいる時に限り，座席までのルート表示，座席からの飲食物の注文，付近の空いてるトイレの案内が可能となる．さらに，試合中にはスタジアム内に設置されたカメラが撮影する選手のリアルタイム映像を視聴できる．

スタジアム内に無線基地局（アクセスポイント，AP）を設置するエリア設計として，従来のマクロセル方式では，多くの利用者がWi-Fiを同時に快適に利用できない問題があった [2]．これは，少ないAP数で省コストである一方，最大同時接続端末数が少ないため，多数の観客が接続すると通信速度が低下するためである．そこでスマートスタジアムは，屋根や壁の上部，観客席の下などのスタジアム内に多数のAPを高密度に設置する，マイクロセル方式を採用している．マイクロセル方式はAP数が多く高コストである一方，スタジアム全体で数万台の同時接続にも対応でき．APごとのパフォーマンスが高く，高水準なサービスが提供可能となる．さらに，APのカバーエリアを狭く限定することにより，AP間の電波干渉を抑制し，動画配信などの大容量コンテンツを快適に利用可能としている．

しかし，観客が持つ携帯端末（以下，利用者端末）が高密度に存在するスタジアムでの通信環境において，低遅延かつ高品質なスポーツ映像を配信するためには，ネットワーク資源の状況を考慮しつつ，端末の利用可能な資源量に適した映像品質を決定する必要がある．既存研究では，映像品質の変動やバッファ占有率に着目することにより，観客の体感品質を向上させているが [3–5]，観客数やサービスの利用状況，利用者端末やスタジアムの特性を事前に把握することは困難であるため，サービス利用時の状況に適応した適切な映像品質を決定できるとは限らない．

そこで本研究では，多段強化学習に基づく映像品質制御手法を提案する．具体的には，利用者端末，エッジサーバおよびクラウドサーバで多段的に強化学習することにより，観客個人およびスタジアムに適応した効率的な映像品質を決定する手法を提案する．本提案により，使用する利用者端末やスタジアムで開催されるコンサート等のイベント種別に依らずに，快適なネットワークサービスの提供が可能になることが期待できる．

本稿では，初期段階の各学習手法とそれらの連携方式について設計する．

[1] 東北大学大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University
[2] 東北大学サイバーサイエンスセンター
Cyberscience Center, Tohoku University
[a)] medio@ci.cc.tohoku.ac.jp
[b)] beto@tohoku.ac.jp
[c)] suganuma@tohoku.ac.jp

2. 関連研究と課題

2.1 関連研究

スポーツイベントにおける映像配信では，スタジアム内に設置した複数台のカメラ映像（以下，自由視点映像）をリアルタイム配信する．このようなサービス形態では，高品質な映像のデータ量に起因するトラヒック低減および遅延低減，ならびに視点切換えによる観客の体感品質の変動が重要となる [6,7]．そのため，クラウドを経由せず，スタジアム内（エッジ）に配信サーバを設置し，映像データの流れをスタジアム内で完結する，地産地消型のデータ流通形態となることが一般的である．

既存研究では，観客が選択した視点映像に関連する視点映像のみ配信することで，ネットワーク帯域の消費量を低減している [3–5]．また，映像品質の変動を対象とした研究では，視点間の映像の関係性に着目することにより，映像品質切替えの頻度と程度を低減している [3]．視点切換えを対象とした研究では，バッファ占有率に着目することで，スムーズな視点切換えと映像再生の停止回数を低減している [4,5]．

2.2 課題

前節で述べたように，既存研究では，視点間の映像の関係性やバッファ占有率に着目することにより，観客の体感品質を向上させている．しかし，観客数やサービスの利用状況はスポーツイベント毎に異なり，利用者端末やスタジアムの特性を事前に把握することは困難であるため，サービス利用時の状況に適応した適切な映像品質を決定できるとは限らない．この問題を解決し，スポーツイベントにおける快適な映像配信サービスを実現するためには，以下の課題に取り組む必要がある．

(C1) 個人化の欠如
利用者端末の性能の考慮が不十分であるため，提供される映像品質が観客の満足する品質ではない可能性がある．

(C2) 公平性の欠如
他の利用者端末に与える影響の考慮が不十分であるため，高品質サービスを低性能端末に提供できない可能性がある．

(C3) 適用性の欠如
他スタジアムや他イベントへの適用可能性の考慮が不十分であるため，効率的な品質制御ができない可能性がある．

3. 多段強化学習に基づく映像品質制御手法

3.1 提案概要

本研究では，多段強化学習に基づくスポーツ映像の品質制御手法を提案する．強化学習を用いることで，利用者端末で利用可能な資源量と映像配信手法から，提供可能な映像品質の関係性を逐次的に学習し，特定の条件下で適切な映像品質を決定可能である [8]．この強化学習を発展させ，利用者端末，配信サーバおよびクラウドサーバで多段的に強化学習することにより，観客個人に特化したサービス品質制御ではなく，スタジアム全体の利用可能な資源に適応した映像品質を効率的に決定する手法を提案する．

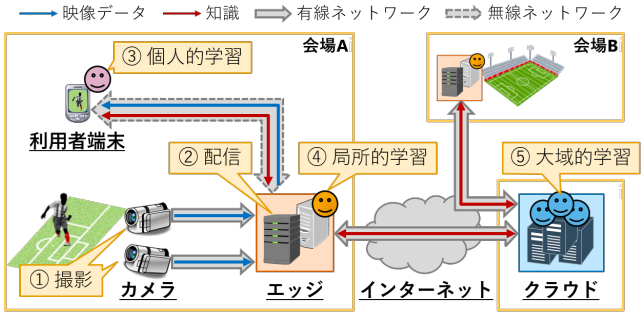

図 1: 提案手法の概要

具体的には，以下の 3 段階の強化学習を導入する．

(F1) 個人的学習
利用者端末の性能に適応し，観客毎の個人化を実現

(F2) 局所的学習
他の利用者端末とスタジアムの特性に適応し，利用者端末間の公平性を充足

(F3) 大域的学習
他のスタジアムやイベントの特性に適応し，スタジアム間の連携を実現

図 1 に提案手法の概要を示す．スタジアムにはカメラの他に，エッジとしてエッジサーバと配信サーバを設置する．エッジとスタジアム外のクラウドにあるクラウドサーバはインターネットを通して接続する．個人的学習，局所的学習，大域的学習は，利用者端末，エッジサーバ，クラウドサーバにてそれぞれ行う．

3.2 提案手法の構成要素

3.2.1 配信サーバ

配信サーバは，複数視点の映像中から利用者端末が要求した品質の映像を配信する．本研究における映像配信では，配信サーバから携帯端末に一方的に映像を配信する従来のストリーミング手法とは異なり，観客が任意の視点を指定し，個人的学習の結果に応じた品質の映像を配信サーバに要求する双方向的な映像配信となる．

そのため，映像の配信方式は Moving Picture Experts Group (MPEG) が標準化した Dynamic Adaptive Streaming over HTTP (MPEG-DASH) [9] とする．MPEG-DASH は，カメラ映像を数秒ごとに分割したセグメントファイルを生成する．セグメントファイルは複数の

品質で圧縮符号化し，配信サーバは，利用者端末からの要求に応じた品質のセグメントファイルを配信する．

3.2.2 利用者端末

個人的学習により決定した映像品質のセグメントファイルを配信サーバに要求し，配信されたセグメントファイルを利用者端末上で再生する．なお，観客は事前に専用のアプリケーションを利用者端末にインストールし，ネットワークはスタジアム内の Wi-Fi を使用するものとする．

個人的学習は，利用者端末に適応した映像品質を学習および決定するため，映像再生中の利用者端末の状態から映像品質を決定する．利用者端末の利用可能な資源量は限られているため，学習手法は，計算負荷が小さい強化学習手法を採用する．学習目標は，低消費資源で再生可能な高品質の映像品質を決定することである．これにより，観客毎の個人化を実現する．

3.2.3 エッジサーバ

エッジサーバで局所的学習により生成したは，利用者端末と共有する．利用者端末では，自身の知識とエッジサーバの知識を統合する．

局所的学習は，特定のスタジアムに適応した知識を生成するため，スタジアム内の各利用者端末から収集した知識群に加えて環境情報を考慮し，各利用者端末に適応した汎用的な知識を生成する．なお，エッジサーバは計算資源が豊富であると想定し，学習手法は，計算負荷の大きい深層強化学習手法を採用する．学習目標は，利用者端末の性能差により生じるネットワーク帯域の割り当ての差を減らすことである．これにより，利用者端末間の公平性を充足する．

3.3 クラウドサーバ

クラウドサーバで大域的学習により生成した知識は，エッジサーバと共有する．エッジサーバでは，自身の知識とクラウドサーバの知識を統合する．

大域的学習は，あらゆるスタジアムに適応した知識を生成するため，各スタジアムのエッジサーバから知識を収集し，膨大な知識から汎用的な知識を獲得する．学習手法は，計算負荷の大きい分散型深層強化学習手法を採用する．学習目標は，汎用的な知識を活用することによる学習の予測精度向上および収束の高速化である．これにより，他のスタジアムやイベントの特性に適応したスタジアム間の連携を実現する．

3.4 動作手順

提案手法を適用したスポーツイベントにおけるライブストリーミングの動作手順を以下に示す．図 1 中の番号と箇条書きの番号は対応している．

① スタジアムに設置した複数台のカメラは，選手やスタジアム全体の俯瞰映像をエッジに送信する．
② 配信サーバは，複数のカメラ映像中から利用者端末からの要求に応じた映像を配信する．
③ 利用者端末は，観客毎の個人化を実現するための個人的学習を行う．具体的には，利用者端末の利用可能な資源量と映像品質の関係性から，低消費資源で再生可能な高品質の映像品質を学習する． 学習した結果に基づき，配信サーバに映像を要求する．
④ エッジサーバは，利用者端末間の公平性を充足するための局所的学習を行う．具体的には，各利用者端末の知識を収集し，利用者端末の消費資源量とスタジアムの特性の関係性から，各利用者端末への帯域の割り当ての差を学習する． エッジサーバの知識は利用者端末に送信し，利用者端末の知識と統合する．映像配信中は①から④を繰り返す．
⑤ 映像配信終了後，クラウドサーバは，他スタジアムと連携するための大域的学習を行う．具体的には，各エッジサーバの知識を収集し，それらの関係性から，汎用的な知識を獲得する． クラウドサーバの知識はエッジサーバに送信し，エッジサーバの知識と統合する．

4. 設計

4.1 個人的学習

個人的学習の学習手法は，Q 学習 [10] を採用する．具体的には，先行研究 [11] に基づき個人的学習をする．Q 学習では，学習の主体となるエージェント（Ag）が報酬 r の期待値であり，行動価値となる Q 値 $Q(s,a)$ を学習していく．Q 値は端末の状態 s と行動 a の組み合わせ数だけ保存する．そのため，状態 s を構成する要素は離散化する必要がある．一般に，s を行，a を列とした表形式で保存することから，s, a, $Q(s,a)$ の組み合わせを Q テーブルと呼ぶ．

Algorithm 1 に Q 学習に基づく個人的学習の疑似コードを示す．初めに，すべての Q 値 $Q(s,a)$ を 0 に初期化する．サービス提供中，Ag は時刻 t における利用者端末の計算機資源 R_{CP_t} とネットワーク資源 R_{NW_t}，映像の視点切換え頻度 fvw から，利用者端末の状態 $s_t \in S$ を観測する．s の全体集合 S を以下に定義する．

$$S = \{s_i \mid s_i = \langle R_{CP_i}, R_{NW_i}, fvw_i, \rangle; i = 1, 2, ..., N_i\} \quad (1)$$

ここで，N_i は利用者端末の総数である．

Ag は観測した s_t に対応した行動 a を方策 π に基づき決定する．本研究では，a の選択は映像品質 $a_t \in A$ を選択することを意味する．a の全体集合 A を以下に定義する．

$$A = \{a_j \mid a_j = \langle rsl, fps, crf \rangle; j = 1, 2, ..., N_j\} \quad (2)$$

ここで，rsl $(\in RSL)$ は映像解像度，fps $(\in FPS)$ はフレームレート，crf $(\in CRF)$ は映像を圧縮符号化する際に指定する画質の程度 Constant Rate Factor (CRF)，N_j は符号化パラメータの総数である．また，π は ε-greedy ア

Algorithm 1 Q 学習に基づく個人的学習

1: Initialize Q values $Q(s,a)$ for all state-action pairs to 0.0
2: **repeat**
3: Get resources (R_{CP_t} and R_{NW_t}) at time t
4: Observe a state $s_t \in S$
5: Determine an encoding parameter $a_t \in A$ according to the policy π
6: Require a video segment with quality a_t to the Distribution Server
7: Receive the segment and a quality degradation $\Delta D_{t+\tau}$ from the Server
8: Play the segment on the display
9: Observe the next state $s_{t+\tau}$
10: Get resources ($\Delta R_{CP_{t+\tau}}$ and $\Delta R_{NW_{t+\tau}}$) at time $t+\tau$
11: Calculate a reward $rwd_{t+\tau}$ according to equation (3)
12: Update the Q-value $Q(s_t,a_t)$ according to equation (4)
13: Transition to next time $t \leftarrow t+\tau$
14: **until** the streaming service is finished.

ルゴリズム基づく決定方法とする．具体的には，確率 ε で受信端末が利用可能な帯域に適応した a_t を，確率 $(1-\varepsilon)$ で Q 値が最大となる s_t に対応した a_t を選択する．

次に，選択した品質 a_t のセグメントを利用者端末で再生し，状態 $s_{t+\tau}$ を再度観測する．さらに，受信端末の資源情報である $\Delta R_{CMP_{t+\tau}}, \Delta R_{NW_{t+\tau}}$ とエンコーダから獲得する圧縮符号化後の映像品質の劣化程度 $\Delta D_{t+\tau}$ から式 (3) に基づき報酬 $r_{t+\tau} \in \mathbb{R}$ を求める．$r_{t+\tau}$ は，a_t が受信端末の状態に与えた影響として Ag に与える実数であり，以下の式で導出する．

$$r_{t+\tau} = -\Delta R_{CP} - \Delta R_{NW} - \Delta D \tag{3}$$

このとき，$\Delta R_{CP}, \Delta R_{NW}$ が増加すると，それぞれ計算機資源，ネットワーク資源の消費量が大きくなるため，r は小さくなる．同様に，ΔD が増加した場合，圧縮符号化後の映像の品質劣化が大きく，サービスの利用者に対して与えるサービス品質が劣化するため，r は小さくなる．

最後に，以下の式で Q 値 $Q_t(=Q(s_t,a_t))$ を更新し，t が $t+\tau$ に遷移する．

$$Q_t \leftarrow Q_t + \alpha\{r_{t+\tau} + \gamma \max_{a' \in A} Q_{t+\tau} - Q_t\} \tag{4}$$

ここで，α は学習率，γ は割引率であり，ともに範囲 (0.0, 1.0] の実数である．

以上の流れをサービス提供中繰り返すことにより，利用者端末やネットワークの状態に適応した高品質な自由視点映像品質を決定する．

4.2 局所的学習

表形式の Q 学習の問題点は，状態 s の要素数が多い場合や，細かく離散化した場合に，Q テーブルの大きさが爆発的に増大することである．膨大な Q テーブルを強化学習するためには非常に多くの試行数が必要となるため，効率的な学習は困難である．

そこで本研究では，この問題点に対処するため，深層強化学習 [12] に基づき局所的学習をする．深層強化学習は深層学習と強化学習を組み合わせた学習手法であり，Q テーブルをニューラルネット化した Q ネットワークとして Q 値の近似関数を得る．しかし，ニューラルネットワークのような非線形関数を用いた Q 関数の近似は一般的に不安定で発散してしまうことが知られている．この原因として，観測する状態 s の要素データ（以降，観測データ）が連続でありデータ間の相関が高いこと等が挙げられる．そのため深層強化学習において安定した学習を実現するための工夫点として，Experience Replay, Fixed Target Q-Network, 報酬の clipping, 誤差関数の選択がある．

Experience Replay は，従来の Q 学習のように 1 ステップごとに Ag がそのステップの経験（experience）を学習するのではなく，メモリに各ステップの内容を保存しておき，メモリから内容をランダムに取り出して（replay）ニューラルネットワークに学習させる手法である．具体的には，Ag の時刻 t の経験 $e_t = (s_t, a_t, r_{t+\tau}, s_{t+\tau})$ をバッファ $B_t = [e_t, ..., e_t]$ にプールする．Q ネットワークの更新時は，B からランダムサンプリングしたミニバッチ $(s,a,r,s') \sim U(B)$ で行う．これにより，観測データ間に生じる相関を除去する．

Fixed Target Q-Network は，主となる main-Q-network とは別に，誤差関数で使用する Q 値を求める target-Q-network を用意する手法である．そして Q 学習で使用する $\max_{a' \in A} Q_{t+\tau} - Q_t$ の値を target-Q-network から求める．このように，main-Q-network を学習する際の $\max_{a' \in A} Q_{t+\tau} - Q_t$ を同じ Q ネットワークから求めるのではなく，少し前の固定しておいた Q ネットワーク（Fixed Target Q-Network）を使用することにより，Q 値間の相関があるために学習が発散，振動しやすくなる問題を緩和する．

報酬の clipping は，Ag が各ステップで得る報酬 r の範囲を限定する手法である．具体的には，-1, 0, 1 のいずれかに固定することにより，学習対象に依らず，同じハイパーパラメータで深層学習を実行可能となる．

誤差関数の選択では，二乗誤差ではなく Huber 関数を使用する．Huber 関数 $H(los)$ は以下の式で表される．

$$H(los) = \begin{cases} \frac{1}{2} l^2 & (|los| \leq \delta) \\ \delta |los| - \frac{1}{2}\delta^2 & (\text{otherwise}) \end{cases} \tag{5}$$

このとき，$\delta = 1$ とすると，誤差 los が範囲 [-1, 1] の間は二乗誤差の値となり，-1 より小さいときや 1 より大きいときには誤差の絶対値をとる．los が大きい場合に二乗誤差を使用すると，誤差関数の出力が大きくなりすぎて学習が安定しにくい問題が発生する．ニューラルネットワークにおいて，los を逆方向に伝搬する誤差逆伝播法により，素子間の結合パラメータを学習していく．

ニューラルネットワークの入力は状態 s_t，出力は全ての映像品質 a_t における Q 値となる．したがって，ニューラルネットワークの入力層の素子数は，状態 s の次元数（$=|S|$）であり，出力層の素子数は，行動 a の次元数（$=|A|$）となる．出力層の素子が出力する値は，Q 値 $Q(s,a)$ となる．そのため，出力層の各素子が出力する Q 値を比較し，一番大きな Q 値を出力した素子に対応する行動を採用する．

Algorithm 2 に Q 学習に基づく局所的学習の疑似コードを示す．初めに，main-Q-network, target-Q-network, リプレイバッファ B を初期化する．サービス提供中，エッジサーバの Ag は観客が持つ利用者端末から Q テーブルを収集する．局所的学習では利用者端末間の公平性を考慮するため，利用者端末から収集した Q テーブルの状態 t に環境情報 ENV を付加して学習する．したがって，ニューラルネットワークに入力する状態 s^+ の全体集合 S^+ を以下に定義する．

$$S^+ = \{s_i^+ \mid s_i^+ = \langle s_i, E_i \rangle; i = 1, 2, ..., N_i\} \tag{6}$$

E は，1AP あたりの観客数 apr と利用者端末が選択可能な映像数 nvw を用いて以下のように表現する．

$$E = \{env_i \mid env_i = \langle apr, nvw \rangle; i = 1, 2, ..., N_i\} \tag{7}$$

Ag は生成した s_t^+ に対応した符号化パラメータ a_t の決定を方策 π に基づき決定する．個人的学習と同様に，π は ε-greedy アルゴリズム基づく決定方法とする．次に，状態 $s_{t+\tau}$ を再度観測し，報酬 $r_{t+\tau}$ を求める．その後，Ag の経験 $e_{i1,t1} = (s_{i1,t1}, a_{i1,t1}, r_{i1,t1}, s_{i1,t1+\tau})$ をリプレイバッファ B に蓄積する．B にある程度経験が蓄積された後，ミニバッチで B からランダムサンプリングした他の利用者端末の経験 $e_{i2,t2} = (s_{i2,t2}, a_{i2,t2}, r_{i2,t2}, s_{i2,t2+\tau})$ から，式 (4) に基づき Q 値の差分 $dif_{i,t}$ を以下の式で求める．

$$dif_{i,t} = (r_{i2,t2+\tau} + \gamma \max_{a' \in A} Q_{t2+\tau}) - Q_{t2} \tag{8}$$

そして，式 (5) を用いて dif から誤差 err を求め，誤差逆伝播法により main-Q-network の結合パラメータを更新する．また，C ステップごとに main-Q-network と target-Q-network を同期するすことを，サービス提供中繰り返す．

これにより，個人的学習の学習効率が向上するための汎用的な知識を生成する．

4.3 個人的学習と局所的学習の連携

個人的学習と局所的学習を連携することで，個人的学習の効率化および公平性の充足を図る．具体的には，局所的学習により学習した main-Q-network の出力層である各 Q 値を個人的学習の Q テーブルに統合する．

Q テーブルの Q 値 Q_{tb} と main-Q-network の Q 値 Q_{nw} を比較し，統合する方式は以下の通りとする．

Algorithm 2 深層強化学習に基づく局所的学習

1: Initialize a main-Q-network Q
2: Initialize a target-Q-network $\hat{Q}$
3: Initialize a replay buffer B to empty
4: **repeat**
5: Get a Q-table from the Receiver $i1$
6: Create a state $s_{i1,t1}^+$ from the Q-table adding the environmental information $E_{i1,t1}$ at time $t1$
7: Determine an encoding parameter $a_{i1,t1}$ according to the policy π
8: Observe $s_{i1,t1+\tau}^+$ and reward $r_{i1,t1+\tau}$
9: Store an experiment $e_{i1,t1}$ in B
10: Sample random minibatch of $e_{i2,t2}$ from B
11: Calculate a difference $dif_{i,t}$ according to equation (8)
12: Calculate an error $err_{i,t}$ according to equation (5)
13: Update the main-Q-network using $los_{i,t}$
14: Reset the target-Q-network to main-Q-network every C steps
15: **until** the streaming service is finished.

$$Q_{tb} = \begin{cases} Q_{tb} + \frac{1}{2}(Q_{nw} - Q_{tb}) & (Q_{tb} \neq Q_{nw}) \\ Q_{nw} & (Q_{tb} = 0.0) \\ Q_{nw} & (\text{otherwise}) \end{cases} \tag{9}$$

これにより，Q テーブルの学習の偏りを抑制し，急激な利用者端末の状態 s の変動にも対応可能とする．さらに，局所的学習により学習した Q 値に対応する行動 a が選択されやすくなるため，公平性の充足を重点においたサービス提供を実現する．

5. おわりに

本研究では，スポーツイベントのライブストリーミングサービスを対象とし，観客が使用する端末が高密度に存在するスタジアムでの通信環境において，低遅延かつ高品質な映像を配信するための手法を提案した．具体的には，ネットワーク資源の状況を考慮しつつ，端末の利用可能な資源量に適応した映像品質を決定するための多段強化学習に基づく映像品質制御手法を提案した．提案手法は，端末，エッジおよびクラウドサーバで達段的に強化学習することにより，観客個人およびスタジアムの環境に適応した効率的な映像品質を決定する．本稿では，各学習手法に焦点を当て，学習の初期段階の学習手法とそれらの連携方式について設計した．

今後は，局所的学習の実装を行い，実験により学習効果を確認する．さらに，大域的学習の詳細や局所的学習との連携方式について検討する．

参考文献

[1] 日本電信電話株式会社：西武ドームにおけるスタジアム エンターテインメントサービスの拡充について，日本電信電話株式会社（オンライン），入手先 ⟨http://www.ntt.co.jp/news2014/1409/140908a.html⟩（参照 2019-12-03）．

[2] 川又英紀：新国立競技場に導入される「高密度 Wi-Fi」のからくり、約 1000 カ所のアクセスポイント用意，日経 xTECH ／日経アーキテクチュア（オンライン），入手先 ⟨https://tech.nikkeibp.co.jp/atcl/nxt/column/18/00933/091200007/⟩（参照 2019-12-03）.

[3] Hamza, A., Ahmadi, H., Almowuena, S. and Hefeeda, M.: QoE-fair Adaptive Streaming of Free-viewpoint Videos over LTE Networks, *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, New York, New York, USA, ACM Press, pp. 161–169 (online), DOI: 10.1145/3126686.3126738 (2017).

[4] Song, S., Kim, Y., Park, Y. S. and Wee, J.: Free-Viewpoint Relationship Description Based Streaming Systems for Arbitrary View Switching, *International Conference on Ubiquitous and Future Networks (ICUFN2018)*, Vol. 2018-July, IEEE, pp. 738–740 (online), DOI: 10.1109/ICUFN.2018.8436845 (2018).

[5] Yao, C., Xiao, J., Zhao, Y. and Ming, A.: Video Streaming Adaptation Strategy for Multi-view Navigation Over DASH, *IEEE Transactions on Broadcasting*, Vol. PP, pp. 1–13 (online), DOI: 10.1109/TBC.2018.2871370 (2018).

[6] Zhao, L. and Chen, Z.: Optimizing Quality of Experience of Free-Viewpoint Video Streaming with Markov Decision Process, *2018 IEEE International Conference on Communications (ICC)*, Vol. 2018-May, IEEE, pp. 1–6 (online), DOI: 10.1109/ICC.2018.8422860 (2018).

[7] Zhang, X., Toni, L., Frossard, P., Zhao, Y. and Lin, C.: Adaptive Streaming in Interactive Multiview Video Systems, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 29, No. 4, pp. 1130–1144 (online), DOI: 10.1109/TCSVT.2018.2819804 (2019).

[8] Gadaleta, M., Chiariotti, F., Rossi, M. and Zanella, A.: D-DASH: A Deep Q-Learning Framework for DASH Video Streaming, *IEEE Transactions on Cognitive Communications and Networking*, Vol. 3, No. 4, pp. 703–718 (online), DOI: 10.1109/TCCN.2017.2755007 (2017).

[9] ISO/IEC: Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats.

[10] Watkins, C. J. C. H. and Dayan, P.: Q-Learning, *Machine Learning*, Vol. 8, No. 3, pp. 279–292 (online), DOI: 10.1007/BF00992698 (1992).

[11] 生出真人, 阿部　亨, 菅沼拓夫：強化学習を用いた MPEG-DASH における映像品質制御手法の実験と評価，第 26 回マルチメディア通信と分散処理ワークショップ (DPSWS2018)，pp. 123–129 (2018).

[12] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. and Hassabis, D.: Human-level control through deep reinforcement learning, *Nature*, Vol. 518, No. 7540, pp. 529–533 (online), DOI: 10.1038/nature14236 (2015).