



Submitted on: 05.07.2021

DOI:10.5281/zenodo.5153235

Using motion capture technologies to assess the degree of similarity for intangible cultural heritage expressions.

Jorge Poveda Yáñez

Performing Arts Department - Faculty of Arts, Universidad Central del Ecuador, Quito, Ecuador.

E-mail address: jmpyanez@gmail.com; jmpoveda@uce.edu.ec.

Benedikte Wallace

RITMO, Department of Informatics, University of Oslo, Oslo, Norway.



Copyright © 2021 by Jorge Poveda Yáñez and Benedikte Wallace. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

Abstract:

After multiple cases of misappropriation of intangible cultural heritage expressions, indigenous communities must still rely on the public outcry to stop third parties from illegitimately exploiting their traditions. Beyond raising awareness of the pervasiveness of these practices, pragmatic tools have to be developed to strengthen the conservation and transmission of traditional cultural expressions in action, such as dances, rituals, and performances, with special attention to the digital environments wherein they circulate.

To defy the apparent contradiction between technology and indigenous culture, the “Movement Similarity Project” explored new technologies of human-movement recognition to understand the possibilities of similarity algorithms and motion capture repositories to overcome the vulnerability in which indigenous dances dwell. These interdisciplinary efforts are raised to protect the interests and livelihoods of indigenous peoples in relation to their intangible cultural heritage, until new parameters to prevent their misappropriation arise on a global scale.

Keywords: Intangible cultural heritage; Traditional cultural expressions; Intellectual Property; Motion capture technology; Human-movement recognition.

Traditional cultural expressions of folklore or intangible cultural heritage that belongs to indigenous peoples dwell in an adverse framework of legal (Torsen & Anderson, 2010) and institutional disparities (Murphy, 2007) that continues to produce controversies and disputes for illegitimate appropriation of songs, dances and culture (Tsosie, 2002: 310). With a renovated vigour, the claims of disrespectful uses of choreographic material by indigenous

peoples have consistently resurged during the digital age (Frankel, 2014; Millán, 2020); and the unprecedented ways of rendering other peoples' dances through new technologies are already resulting in legal disputes of authorship and ownership, as seen with the videogame "Fortnite" (Goslin, 2019). Considering this convoluted background, and since the more than twenty years of negotiations at the World Intellectual Property Organization (WIPO, 2019) are still yet to delimit a more egalitarian field for the circulation of traditional cultural expressions, the authors of this paper started exploring alternatives to confront the issue through human-movement recognition technologies. The common factor between the appropriation of the *Ka Mate Haka* dance of the Māori people by The New Zealand Rugby team (Frankel, 2014); of the Hopi ritual dances for the commercial benefit of non-Hopi people (Torsen & Anderson, 2010: 76); the Sierra dance of Peru (WIPO, 2002: 32) or the *Choike Purrún* Dance of the Mapuche people (Millán, 2020) is that their traditional bearers needed to convince audiences and/or authorities, that the dance in dispute, was indeed *their* dance. In order to support and assess the claims of misappropriation of expressions of intangible cultural heritage, the "Movement Similarity Project" carried a series of in-lab experimentations, using motion capture technology, to quantitatively measure the degree of resemblance between two different performances of the same dance. In the first section of this paper, a reflection about the relevance of this project for the protection of the intangible cultural heritage of indigenous peoples is raised. Section two introduces the scope of the "Movement Similarity Project", section three describes the data collection process, section four includes the details of the movement analysis techniques employed, section five includes a discussion of the findings and section six foregrounds the preliminary conclusions.

The alliance between indigenous peoples' interests and new technological developments

The multimodal nature of dance and therefore of the different types of misappropriation of indigenous dances that can be done by third parties require specific metrics and systems of appraisal. Indeed, according to the case, the appropriation of indigenous dances alternatively involves the names, the symbols, words referring to myths, the costumes, the songs or lyrics that pertain to an indigenous community. Regardless of the layer of dance that is allegedly being misappropriated, such situations will always have to be assessed and measured by decision-makers, if sanctions and remedies are expected. 'Cultural heritage rights claimants will need to solicit and rely on judicial deference (...) The challenge for Courts will be to differentiate between legitimate and illegitimate statements of such beliefs and practices' (Paterson and Karjala, 2003: 669). Third parties' appropriation of an indigenous dance can be focused solely on the use of the steps, without using any referential words, legends or symbols, rendering the situation less visible to be addressed and controverted. What becomes relevant is that regardless of the system of protection invoked, there is an underlying need for tools to empirically and quantitatively establish if there has been such a thing as an appropriation of an indigenous dance. 'In all cases, the plaintiff will have to prove that the defendant copied from the copyrighted material in a way that is qualitatively or quantitatively substantial' (Burri, 2018: 21). In this sense, to partially address the urgent need expressed by indigenous peoples about the 'prevention of insulting, derogatory and/or culturally and spiritually offensive uses of traditional cultural expressions' (WIPO, 2018: 8), a combination of digitisation tools and methods, such as motion capture, can be employed.

New technologies of human movement recognition including motion capture and video-visualization techniques, can be suitable to provide quantitative reports for Courts to examine disputed dances. Moreover, similarity algorithms can be a strategic asset for plaintiffs to prove that there has been a kinetic material illegitimately appropriated by third parties.

Courthouses will not settle with just an apparent resemblance, and therefore it is a common practice to require the expert opinion of professionals in the specific area of discussion to establish a significant resemblance. In this sense, the “Movement Similarity Project” was designed with the premise of exploring different computer-based metrics that could shed a light on the resemblance between two different performances of the same dance.

The “Movement Similarity Project”

Measuring and comparative tools should focus on more than the mere sequence of key postures that the bodies go through and also consider the way in which these choreographic elements are being performed. For the “Movement Similarity Project”, not only the steps involved in the recorded choreography were considered, but also the stylistic traits that separate one performance from another, despite how similar the underlying choreography may be. For this reason, the “Movement Similarity Project” explored several methods for measuring movement – some focused on comparing qualitative aspects, others aimed towards measuring content similarity. This project involved recording full-body motion capture of 10 participants dancing the same choreography while being prompted to dance with varying degrees of enthusiasm. Thereby, movement quality aspects can be contrasted while keeping the underlying content of the movement identical. For the project presented here, an optical motion capture system with passive markers has been used. The motion tracking system employed consisted of 12 Oqus 300 and 400 motion tracking cameras which use infrared light to capture the reflective markers at a frame rate of 250Hz. This system ensures precise recordings from which features of interest could be calculated. Details concerning the experiment design and the data collection process for this project are presented in the following section.

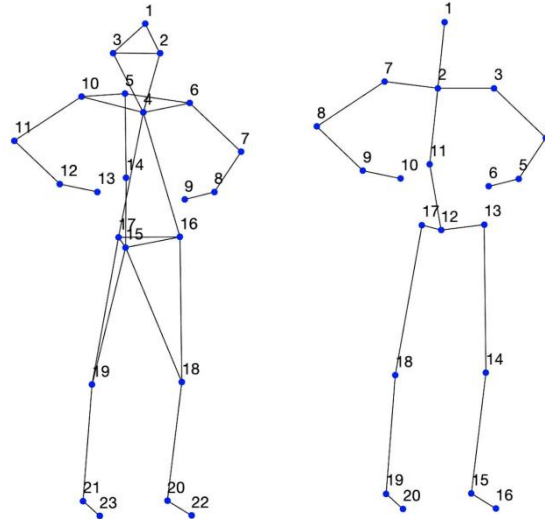


Figure 1 Marker placement (a) and corresponding body segments with joints (b)

Data Collection

An important consideration when collecting the data was making sure that the choreography was easy to learn so that subjects with little to no previous dance experience could still participate. For this reason, the *Macarena* dance was chosen. Most people are familiar with the steps involved and only needed a few minutes to learn the dance. To keep performances on time, a metronome was played at 110 BPM for participants to dance along to. A single

performance lasted 30 seconds. Participants were recruited from individuals connected to the University of Oslo and the motion capture lab of the RITMO Centre. Ten subjects were invited to participate in the recording sessions for this project, 3 of them being female. The participants were dressed in motion capture suits which allowed the reflective markers to be attached. Each participant wore 23 markers as shown in Figure 1a, along with a 24th marker which is used for reference during post-processing and disregarded for analytical purposes. Participants were informed that they would be performing several repetitions of the *Macarena* accompanied by a metronome. They then spent 3 to 5 minutes learning the choreography before the recording began. Markers were placed on the participants according to the CMU marker placement guide 2 (see Figure 1) which uses bony landmarks on the body to position markers consistently across participants. During motion recording, each participant performed the *Macarena* in three different modes, *normal*, *deadpan* and *enthusiastic*.

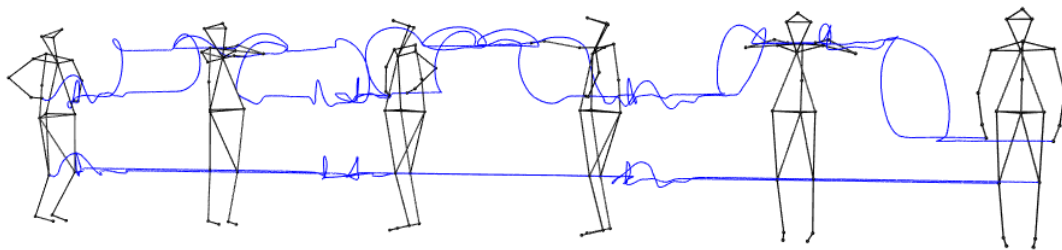


Figure 2 Keyframe plot showing the movement of a performer over time in “deadpan” mode.



Figure 3 Keyframe plot showing the movement of a performer over time in “enthusiastic” mode.

Each participant performed 4 repetitions of the choreography without further instruction (normal mode), 3 repetitions under instructions to make no unnecessary movements (deadpan mode), and 3 repetitions with instructions to dance as if they “really loved it” (enthusiastic mode). Thereby, we collected 10 recordings for each of the 10 participants resulting in 100 recordings in total. As with all optical motion tracking systems, this system is sensitive to marker occlusion and therefore, some post-processing steps were performed before the data could be analysed. These steps include labelling markers according to their position on the body using the Qualisys Track Manager 2019.3 recording software and using interpolation to fill gaps in the trajectories of markers that suffered from dropout. To prepare the recordings for movement analysis, all examples were trimmed and synced to match the starting frame of a single example. This is done using windowed cross-correlation for all markers.

Comparing different repetitions of the same dance

In this work, the definition of movement similarity is divided into two parts: content similarity and quality similarity. Both are described in further detail below, together with the methods used to analyse each aspect. The various methods of analysis are performed using the MoCap Toolbox.

Movement content similarity

The ability to recognise specific gestures is useful in both human-computer interaction, human-robot interaction and information retrieval tasks. When recognising certain motion sequences, it is often advantageous for a system to be invariant to the way a movement is performed as the personal characteristics of users may differ greatly and a good system would need to be robust against certain motion features such as temporal, spatial and stylistic differences. When the target gesture is known one approach is to create a robust description or template of the movement with which you can compare other examples. Although motion templates could work well for the *Macarena* recordings in which the participants are all performing the same choreography, this approach is not necessarily as useful for comparing examples that may contain different forms of dance. Creating templates for each possible pose or movement sequence in the world of dance would require an extensive, if not an insurmountable amount of work.

Here, two more general approaches were implemented. Firstly, the Euclidean distance between the marker trajectories of each performance is measured. To ensure the distance scores are invariant to the spatial location of the participant the data is first normalised in reference to each recording's respective root marker, thereby creating a local coordinate system fixed within the body. The second content similarity metric used involves measuring the differences between sequences of joint angles from several body segments. When using joint angles instead of marker positions in a local or global coordinate system the results are invariant to both the spatial position of the body and variation in limb length. By treating the angle between different segments at each frame as a sequence the recordings can be compared using dynamic time warping, resulting in a collection of pairwise dissimilarity scores. Dynamic time warping (DTW) is a common approach to time series comparison used in various signal processing tasks. Although two sequences may be different in regards to their Euclidean distance, they may still share an overall shape that does not necessarily line up with regards to timing. By stretching or compressing one or both time series the DTW algorithm finds the minimum distances between each point in the two sequences. The sum of these minimum distances produces a score that is interpreted as the amount of dissimilarity. Identical time series will have a DTW score of zero. Although the aforementioned methods facilitate the clustering of similar examples without pre-existing knowledge of what motions the examples contain, using DTW can be a slow and computationally heavy operation and as such these methods are often too costly to be implemented in certain systems.

In the case of information retrieval, for example, the need for fast, content-based indexing of motion clips imposes the need for high-level features such as Boolean features of geometric relations between body segments. In recent years, there has been an increase in machine learning (ML) methods for motion recognition and generation. Holden et al. (2016) use convolutional autoencoders to synthesise the movement of an animated character along trajectories decided by the user. Abidine et al. (2018) perform more efficient human activity recognition (HAR) using weighted support vector machines, and Hassan et al. (2018) perform

HAR using deep neural networks trained on wearable body sensor data. Although trained ML models can be relatively compact and are often inexpensive at run-time, training the models often require large amounts of data and processing power. Using ML for clustering or classifying movement also requires labelled data. This can be problematic as there are limited labelled datasets for dance moves. For this project, joint angle sequences and Euclidean distance are used as the two content similarity metrics.

Movement quality similarity

Movement qualities are often associated with emotional expression. Therefore, being able to recognise these qualities or transpose a movement quality from one movement sequence to another is useful in the world of computer animation and also in human-robot interaction. In order to examine and compare movement qualities, we used an approach loosely based on Laban Movement Analysis (LMA) (1988). While LMA has traditionally been used in dance education and analysis, recent works argue that the use of systems to recognise Laban Movement Efforts helps to improve human-computer interaction (Alaoui et al., 2017). The Laban effort factors are weight, space, time and flow. Each of the effort factors is considered as a continuum between two opposites as described in the following paragraphs.

- Sustained or Sudden

Sudden and sustained movements are characterised according to their duration in time and are associated with motions such as glide (sustained) versus punch (sudden). We measure how sudden or sustained a movement is by considering the amount of acceleration as it changes over time. How similar two examples are with respect to this movement feature can then be measured by comparing the acceleration curves for both examples using DTW. When considering this feature, the right-hand marker is used (Marker 9 in figure 1). In the *Macarena*, most of the movement happens in the arms and as left and right arms mirror each other in this choreography, the choice between left and right arm has been made arbitrarily.

- Direct or Flexible

This feature relates to the transition between what we might think of as poses. When a body segment moves from point A to point B, does it do so in a direct way, or is the path more flexible? When measuring where on the continuum an example lies we calculate the cumulative distance travelled. To compare two examples we calculate and compare the cumulative distance travelled by the left and right-hand markers during consecutive frames. If the participants' movements are flexible we expect to see a higher value for distance travelled, while if the transitions are more direct the value should be lower.

- Gentle or Firm

In LMA, gentle movements are associated with actions such as a flick or a glide, while firm movements are associated with motions like a slash or a punch. To compare the similarity of gentle or firmness between two performances, velocity over time is used. The velocity of the right-hand marker is calculated for each performance and the resulting velocity curves are compared using DTW.

- Jerk or Flow

We can think of jerk as the number of abrupt changes that are observed in a sequence of movement. Jerk is measured using the derivative of acceleration. This feature is measured by

calculating the third time derivative for the right-hand marker and comparing the distance between two performances calculated using DTW.

Discussion

After engaging with motion capture and similarity algorithms, as a means to develop tools that could potentially support indigenous peoples' claims of misappropriation of their dances, we dwell on the importance of not only digitising expressions exclusively when a controversy arises but in a prospective and preventive way too. A motion capture archive of indigenous cultural expressions should aim to operate always at a transnational level to build cross-cultural networks between sub-national entities that can be vigilant of each other's work and creations. The project of a digital repository of indigenous dances is inherently deconstructive of the intention to keep framing the contemporary world under the binary of modern versus traditional. Instead, such a project surpasses these labels by intersecting a modernity permeated by the traditional, rendering visible the mutual influx of these fields instead of polarising them. Then again, it becomes relevant to reconsider the imperative of designing systems that indigenous peoples can use, immerse and benefit from. It is essential to distinguish such participatory archives from the conventional ones, where "indigenous voices have had little impact in shaping the information architectures that underlie how cultural heritage is organised and disseminated online (...) Such attempts at cultural heritage documentation may have little relevance to a local community; worse, it may be seen as an act of appropriation, even if motivated by a desire to help" (Kansa, 2009, cited in Burri, 2018: 16).

The dilemma for indigenous peoples has been presented sometimes as a dichotomy between commodifying their own practices or by renouncing this, to open the door for third parties to freely profit from their exploitation. This paper contends that such binarism is a pervasive illusion that does not fully consider how a well-managed system of documentation, through new technologies, can assist indigenous populations in their pretensions to either obtain property rights or merely prevent others from so doing. In this sense, indigenous peoples will finally be seen not only as groups that either resist or surrender to the future but rather as builders of their own future(s). What this fundamentally suggests is a delicate balance between openness and enclosure, between the publication of materials and secrecy. In other words, to evolve from the current state of disrespectful exploitation to a cross-cultural awareness based on mutual consideration. 'Native culture must not only be protected by (...) laws but must also be allowed to flourish. Thus, too strict an application could destroy (through stagnation) the very culture the law was designed to protect' (Long, 1998: 151). Instead of censoring and dismissing any exchange between cultures, the intention of more fair conditions of cross-cultural learning has to make up for the asymmetries between groups.

Motion capture allows for extremely precise recordings of motion, this could ensure that reuse of performances can be uncovered using simple difference metrics such as Euclidean distance between markers. However, from the aforementioned project, we find that several metrics have a significant impact on similarity measures. Across the movement quality features, the largest variation is found between examples in enthusiastic mode compared to the other modes. There is also more within-group variation for examples in enthusiastic mode compared to normal or deadpan performances. When examining the content similarity features it becomes clear that movement quality features, as well as person-specific characteristics, also affect these measurements.

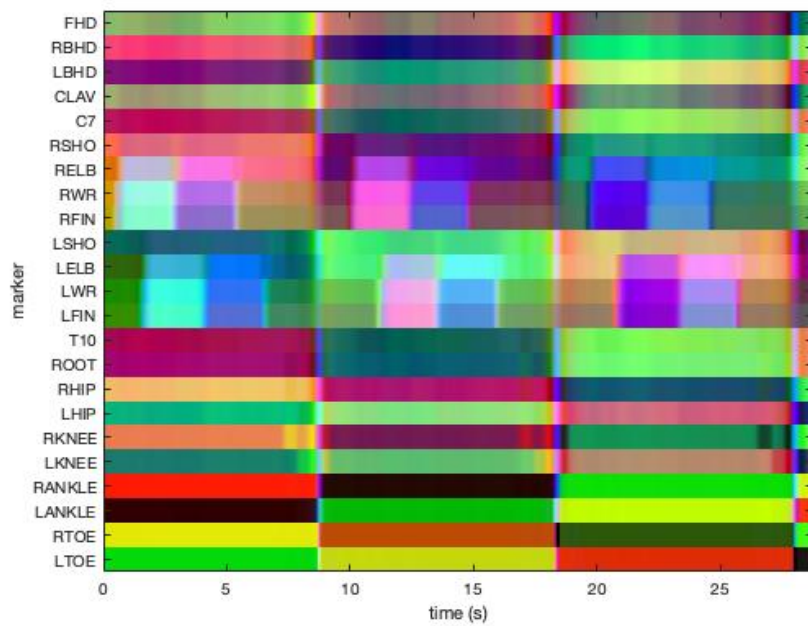


Figure 4 MocapGram of the “deadpan” quality with the specification of each body marker

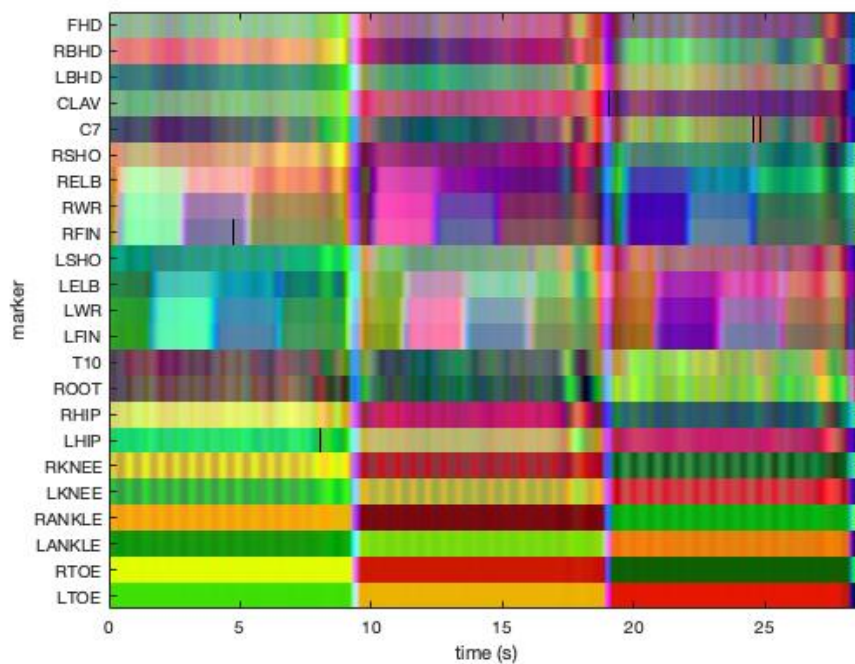


Figure 5 MocapGram of the “enthusiastic” quality with the specification of each body marker

Visual aids such as the mocapgram offer an initial overview to compare performances at a glance. The mocapgrams map the x, y and z coordinates of each marker to the RGB color space creating a visual interpretation in which the content similarity between examples becomes clear. Certain qualitative aspects are also visible, such as the bouncing of the knees and slight timing differences caused by more expressive movements common in enthusiastic performances are reflected in the more diffuse color blocks related to the arm markers. This

visual representation shows how the overall structure or content of the movement can be somewhat lost when comparing examples based on the distances between markers or joint angles, as these expressive aspects may distort the measurements.

Conclusions

Cases of misappropriation of indigenous dances that imply the utilisation of choreographic elements posit a problem for decision-makers to assess the degree of resemblance. It is the contention of this work that such resemblance can be revealed by new technologies of human movement recognition. Motion capture and similarity algorithms were theoretically and practically explored. Satisfactory results were obtained through the “Movement-Similarity Project” as the degree of similarity observed, at first sight, was coincidental with computer measurements. This represents an asset since decision-makers require quantitative and empiric reports to settle disputes of any kind, and the issues surrounding the misappropriation of dance are no exception.

Six different methods were presented, two methods for measuring content similarity and four for measuring movement quality similarity. The results showed how each method was able to derive a partitioning of the data. While the methods described here were able to cluster the data in ways that shed light on the underlying differences between examples, we found the best performance when heuristic choices rooted in knowledge of the recordings are applied. To achieve a more general movement similarity metric, based on movement qualities or content, a combination of several in-depth approaches would be needed as well as participation from individuals with underlying insights into the performance style. In this sense, the participation and involvement of expert practitioners of each intangible cultural heritage traditional are relevant, both for moral and technical reasons.

The isolation of indigenous cultures and a total overlook of digitising and documenting strategies make it easier for third parties to appropriate traditional knowledge. The delicate balance that must be achieved between openness and secrecy can be achieved through smart data-management systems that could document and record traditions while making them available according to the interests and considerations of the bearers. Valid concerns revolving around cultural assimilation and commodification should not lead to the flawed approach of considering that indigenous culture and technologies are at odds because such approach works against the moral and economic interests of indigenous peoples in the long term. It is the contention of this paper that cultural managers, academia and researchers can employ the described metrics for the documentation, digitisation and curation that is required for the protection of indigenous dances, in and out of Courts, always with the guidance and participation of indigenous communities at the centre.

Acknowledgments

The authors would like to thank Kristian Nymo for his guidance and helpful discussions. This work was partially supported by the Research Council of Norway through its Centres of Excellence scheme, project number 262762 as well as the Erasmus Mundus Joint Master Degree (EMJMD) Choreomundus - International Master in Dance Knowledge, Practice and Heritage.

References

- Abidine, B. M., Fergani, L., Fergani, B., & Oussalah, M. (2018). The joint use of sequence features combination and modified weighted SVM for improving daily activity recognition. *Pattern analysis and applications*, 21(1), 119–138.
- Alaoui, S.F., Francoise, J., Schiphorst, T., Studd, K. & Bevilacqua, F. (2017). Seeing, sensing and recognizing laban movement qualities. *Proceedings of the CHI Conference on Human Factors in Computing Systems, USA*, 4009–4020.
- Burri, M. (2018). *Cultural Heritage and Intellectual Property*. Rochester.
- Frankel, S. (2014). “Ka Mate Ka Mate” and the protection of traditional knowledge. In A. Crawford & L. Lincoln (Eds.), *Intellectual Property at the Edge* (pp. 193-214). Cambridge University Press.
- Goslin, A. (2019, March 11). *Fortnite dance lawsuits dismissed after new supreme court ruling*. Polygon. <https://www.polygon.com/fortnite/2019/3/11/18260142/fortnite-dance-lawsuits-dismissed>.
- Holden, D., Saito, J., & Komura, T. (2016). A deep learning framework for character motion synthesis and editing. *ACM Transactions on Graphics (TOG)*, 35(4), 138-149.
- Laban, R. (1988). The movement and the body, in R. Laban and L. Ullmann (Eds.), *The mastery of movement* (pp. 95-148). Plymouth: Macdonald and Evans.
- Long, D.E. (1998). The impact of foreign investment on indigenous culture: An intellectual property perspective. *North Carolina Journal of International Law and Commercial Regulation*, 23(2), 229-240.
- Hassan, M., Huda, S., Uddin, M. Z., Almogren, A., & Alrubaian, M. (2018). Human activity recognition from body sensor data using deep learning. *Journal of medical systems*, 42(6), 99-107.
- Millán, M. (2020, October 31). *Pueblos Originarios. La lidereza mapuche, Moira Millán denuncia a una coreógrafa Europea por extractivismo cultural* [Orinary Nations. Mapuche leader, Moira Millan denounces cultural appropriation by European choreographer]. Quinto Poder Cono Urbano. <https://quintopoderconurbano.com/inicio/2020/02/24/>.
- Murphy, J. (2007) *The People have Never Stopped Dancing: Native American Modern Dance Histories*. University of Minnesota Press.
- Paterson, R. and Karjala, D. (2003). Looking beyond intellectual property in resolving protection of the intangible cultural heritage of indigenous peoples. *Cardozo Journal of International and Comparative Law*, 11(3), 633-671.
- Torsen, M. & Anderson J. (2010). *Intellectual property and the safeguarding of traditional cultures: Legal Issues and Practical Options for Museums, Libraries and Archives*. WIPO publishing.

Tsosie, R. (2002). Reclaiming native stories: An essay on cultural appropriation and cultural rights. *Arizona State Law Journal*, 34(1), 299-313.

WIPO/Intergovernmental Committee on Genetic Resources, Traditional Knowledge and Folklore (2018, May 24). *The Protection of Traditional Cultural Expressions: Updated Draft Gap Analysis*. https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=426450.

WIPO/Intergovernmental Committee on Genetic Resources, Traditional Knowledge and Folklore. (2019, September 14). *Traditional cultural expressions: a discussion paper*. https://www.wipo.int/edocs/mdocs/tk/en/wipo_grtkf_ic_40/wipo_grtkf_ic_40_13_rev.pdf.

WIPO (2002, May 2). *Consolidated analysis of the legal protection of traditional cultural expressions/expressions of folklore*. https://www.wipo.int/edocs/pubdocs/en/tk/785/wipo_pub_785.pdf.