

## **Automatyzacja procesu przenoszenia mimiki aktora na docelową siatkę twarzy**

Rozwój technologii towarzyszących animacjom filmowym oraz grom komputerowym sprawił, że zaspokojenie oczekiwań odbiorców tych mediów odnośnie realizmu prezentowanych scen jest zagadnieniem tak trudnym jak i pracochłonnym. Przygotowywanie animacji jest procesem często opartym o wrażenia estetyczne animatora, jego wiedzę i doświadczenie. Choć w przypadku animowania obiektów, które nie mają swoich odpowiedników w rzeczywistości (np. stworzenia pochodzące z literatury fantasy) proces ten zależy w dużej mierze od wyobrażeń autora, to w przypadku obiektów reprezentujących swoje rzeczywiste odpowiedniki (sylwetka człowieka, twarz, dłonie) możliwym jest uzyskanie wiedzy na temat rzeczywistych procesów, która posłuży stworzeniu realistycznie wyglądającej animacji komputerowej. Rozwój technik i oprogramowania motion capture umożliwił niemal bezpośrednio przenoszenie ruchu aktora na jego trójwymiarowy model. Proste przekształcenia opartej o ludzki szkielet hierarchii pozwalają na łatwe odwzorowanie ruchu w realistyczny sposób, dzięki czemu możliwa jest automatyzacja pracy animatora w zakresie motion capture całego ciała. W przypadku miękkich, złożonych powierzchni, jak ludzka twarz, automatyzacja tego procesu jest zdecydowanie bardziej skomplikowanym zagadnieniem. Z tego powodu standardem w przemyśle gier komputerowych, symulacji, filmów animowanych i efektów specjalnych jest zastosowanie performance capture jedynie, jako inspiracji lub jako wstępnego przekształcenia zestawu punktów na twarzy, które potem poprawiane jest i uzupełniane przez animatora. Zautomatyzowanie procesu estymacji powierzchni twarzy w zależności od przyjętej mimiki określonej w procesie performance capture mogłoby znacznie skrócić czas pracy animatora, ograniczając go jedynie do końcowych poprawek, lub też nawet zlikwidować zapotrzebowanie na manualną ingerencję w taki proces. Ze względu na wymagane umiejętności, doświadczenie oraz żmudność modyfikacji wysokorozdzielczej siatki w sposób pozbawiony automatyzacji, jej rozwinięcie daje możliwość zredukowania znacznych kosztów związanych z wytworzeniem produktu zawierającego animowane twarze, a tym samym przekierowanie tychże środków w obszary związane bardziej z treścią niż formą danej wizualizacji.

W ramach projektu doktorskiego przewidywana jest realizacja badań na podstawie dwóch rodzajów zestawów danych, odpowiednio związanych z markerową i bezmarkerową techniką performance capture. W ramach markerowej techniki, model będący przedmiotem animacji będzie uzyskany przy użyciu skanowania 3D (metodą światła strukturalnego poprzez skaner 3dMDface i przy użyciu stereowizji i wielu zsynchronizowanych aparatów fotograficznych), natomiast dane o mimice aktora, które posłużą stworzeniu animacji zostaną uzyskane przez oparte o światło bliskie podczerwieni kamery systemu VICON. Technika bezmarkerowa oparta będzie o dane pochodzące z systemu sześciu kamer światła widzialnego. Obie te techniki stanowią różne podejście do zagadnienia performance capture i w różny sposób prezentują trudności związane z automatyzacją procesu tworzenia animacji.

Podstawowym problemem w przypadku techniki markerowej jest to, iż animowany model nie może pochodzić bezpośrednio z nagrania, które opisuje mimikę. Musi być on stworzony niezależnie. Nawet w przypadku, w którym osoba zostanie zeskanowana dostarczając realistyczny, wysokorozdzielczy model twarzy, a następnie ta sama osoba posłuży jako aktor dostarczający danych o mimice, zmiany występujące pomiędzy akwizycją twarzy i mimiki, a także brak możliwości uzyskania identycznego wyrazu twarzy prowadzą do nieodpowiedniości modeli. Istnieje więc konieczność przeniesienia danych do innej przestrzeni, określonej przez animowany model. To przeniesienie nie jest prostym przekształceniem jednolitym dla całej przestrzeni. W rzeczywistości, twarz zmienia się w zależności od struktury kości i mięśni pod wpływem takich czynników jak zmęczenie czy stres. Prawidłowe przeniesienie danych uzyskanych w procesie performance capture na model twarzy musi opierać się więc na położeniu punktów charakterystycznych, wykrywalnych w sposób automatyczny

*Damian Pęty*

na modelu twarzy za pomocą krzywizn jego powierzchni i cech antropometrycznych. Podobnie, markery umieszczone muszą być w miejscach, które będą takim charakterystycznym punktem odpowiadały. Wszelkie przesunięcia markerów po przeskalowaniu mogą być więc zastosowane do punktów charakterystycznych na animowanym modelu, co pozwala na przeniesienie mimiki w skali kilkudziesięciu punktów. W przypadku wysokorozdzielczego modelu pozostawia to jednak dziesiątki tysięcy pozostałych wierzchołków modelu w pozycji neutralnej. W ramach projektu doktorskiego planuje się opracowanie metod estymacji powierzchni międzymarkerowej w sposób pozwalający możliwie realistycznie odwzorować mimikę twarzy również w miejscach, które nie są bezpośrednio związane z leżącymi na nich markerami.

Zagadnienie bezmarkerowego performance capture likwiduje problem konieczności estymacji zachowania powierzchni pomiędzy punktami pomiaru – cała powierzchnia twarzy jest próbkowana w rozdzielczości wystarczającej na sterowanie każdym wierzchołkiem wysokorozdzielczego modelu. Problemem jest jednak prawidłowe zbudowanie modelu w oparciu o obraz z kilku kamer i następnie śledzenie każdego wierzchołka zbudowanego modelu na podstawie zmian koloru pikseli w obrazach. Zmiana warunków otoczenia; jasności, koloru i kierunku oświetlenia czy temperatury, jak i same emocje wymuszające zmianę w dystrybucji krwi i tym samym odcień skóry utrudniają śledzenie poszczególnych fragmentów skóry. Opracowywany w ramach pracy doktorskiej algorytm śledzenia musi więc działać na podstawie modelowania sąsiedztwa każdego fragmentu skóry w sposób niezależny od absolutnej wartości barwy.

Opracowanie powyższych zagadnień zarówno w kontekście bezmarkerowego jak i markerowego performance capture pozwoli na użycie obserwacji pozyskanych w ramach tworzenia algorytmów dla jednej techniki w drugiej. Określone w ten sposób zostaną ograniczenia związane z użyciem systemu markerowego, gdzie ilość próbek jest stosunkowo niewielka, w porównaniu do systemu wizyjnego, gdzie śledzenie jest znacznie bardziej podatne na błędy. Dzięki opracowaniu algorytmów retargetingu dla markerowego podejścia, możliwym stanie się zastosowanie podobnych, lecz rozszerzonych algorytmów w oparciu o większą ilość próbek pozyskanych w systemie bezmarkerowym, przy jednoczesnym oparciu o punkty antropometryczne.

**Damian Pęszor**

