



Tytuł: **Percepcja głębi**

Autor: doc. dr hab. Robert Piłat; rpilat@ifispan.waw.pl

Źródło: <http://kognitywistyka.prv.pl/>; mcmarcus@poczta.onet.pl

0. Wstęp

Przedmiotem poniższych rozważań jest głębia pola widzenia. Wyjaśnienie zjawiska głębi oznaczałoby udzielenie odpowiedzi na dwa pytania: (1) Jaki jest stosunek widzianej trójwymiarowej przestrzeni do przestrzeni fizycznej?; (2) Jaka jest funkcja a w szczególności funkcja poznawcza naszej percepcji głębi? (3) Jaki jest neurofizjologiczny mechanizm widzenia głębi i jak się on ma do innych mechanizmów zmysłowych dostarczających nam informacji o przestrzeni fizycznej? Analizując, z konieczności fragmentarycznie, te pytania, wykorzystam dwa punkty widzenia, fenomenologiczny i kognitywistyczny. W kolejnych paragrafach nawiążę do: (1) klasycznego problemu Molineaux dyskutowanego przez Locke'a i Leibniza, a współcześnie analizowanego przez Evansa (1985), Millikan (1991); (2) trudności, jakie napotyka teoria widzenia Marra (1982) w wyjaśnieniu widzenia dwuocznego; (3) fenomenologicznych opisów głębi, jakie znajdujemy u Merleau-Ponty'ego w *Fenomenologii percepcji* i eseju malarstwie Cezanne'a. Moje wnioski podtrzymują i rozwijają uwagi Merleau-Ponty'ego i Evansa o tym, że koniecznym składnikiem widzenia głębi jest pewna wpojona (zautomatyzowana) sprawność psychomotoryczna. Umożliwia ona znajdowanie parametrów dla wirtualnych przestrzeni niezbędnych do zdobycia informacji o przestrzeni realnej na podstawie własności światła rejestrowanych przez siatkówki.



1. Problem Molineaux

Dociekliwy szkocki uczony Molineaux zwrócił się do Johna Locke'a z pytaniem, jaki byłby rezultat następującego eksperymentu myślowego: "Wystawmy sobie, iż niewidomy od urodzenia, obecnie człowiek dorosły, nauczył się odróżnić dotykiem sześcian od kuli, wyrobione z tego samego metalu i w przybliżeniu tych samych rozmiarów, tak że dotykając jednego lub drugiego może powiedzieć, które jest sześcianem a które kulą. Przypuśćmy dalej, że sześcian i kulę ustawiono na stole i że stojący opodal niewidomy przejrzał; zapytuję, czy za pomocą odzyskanego wzroku, zanim dotknie tych przedmiotów, będzie mógł je rozpoznać i powiedzieć, który z nich jest kulą a który sześcianem? (...) Nie; bo chociaż wiadomo mu z doświadczenia, jak działa na jego dotyk kula, a jak sześcian, to jednak brak mu jeszcze doświadczenia, które by go pouczyło, jak to, co taka a tak działa na jego dotyk, w każdym przypadku działa na jego wzrok (...) Zgadzam się ze zdaniem tego, tak poważnie myślącego korespondenta (...) myślę jak on, że były niewidomy nie będzie zdolny natychmiast po zobaczeniu tych ciał powiedzieć z całą pewnością, co jest kulą a co sześcianem" (Locke 1955, t. 1, s. 182)

Trzeba powiedzieć, że pytanie zadane było starannie, jako że Molineaux zdawał sobie sprawę, że gdyby prezentowane przedmioty były znacząco różnej wielkości, to odpowiedź zostałaby udzielona przed odwołaniem się do tego, jak wielką część pola percepcji byłaby zajęta przez oba ciała, tu zaś mogłoby ingerować przeszłe doświadczenie byłego niewidomego, który mógłby być bardziej oswojony – powiedzmy – z dużymi sześcianami niż z dużymi kulami. Stosowałby wówczas pewną empiryczną generalizację, a następnie dedukcję, nie rozpoznawałby zaś cech przestrzennych odróżniających sześcian od kuli. Analogiczną funkcję w eksperymencie myślowym pełni wymaganie, by sześcian i kula były z tego samego materiału. Własności związane z odbijaniem światła, jak faktura, mogłyby stać się podstawą generalizacji i dedukcji, nie zaś rozpoznania brył na podstawie widzenia.

Trzeba pamiętać, że wbrew pozorom, problem Molineaux jest eksperymentem myślowym a nie pytaniem empirycznym. Wydawałoby się, że nic prostszego, jak znaleźć opisany przypadek – jakkolwiek rzadko by występował – i sprawdzić. Wynik obserwacji nie dalby nam jednak pełnej odpowiedzi. Nastąpiłaby ona wówczas, gdybyśmy dowiedzieli się w jaki sposób wyleczony niewidomy odróżnił bryły, lub dlaczego nie mógł ich odróżnić. Tu oczywiście sama obserwacja nie wystarczy; trzeba wysunąć teorię widzenia własności przestrzennych przedmiotów, czyli odpowiedzieć na pytanie: Jakie własności naszego aparatu wzrokowego pozwalają nam na rozpoznawanie (a zatem kategoryzowanie) własności przestrzennych.

Locke odpowiedział na zadane przez Molineaux pytanie negatywnie, kierując się swoim ogólnym stanowiskiem epistemologicznym. W myśl Locke'owskiego empiryzmu, składnikiem percepcji świata są wrażenia zmysłowe, które muszą być związane przyczynowo z przedmiotami widzianymi. Ponieważ zakres przyczyn fizycznych oddziałujących na poszczególne zmysły jest inny (fale akustyczne, fale świetlne, cząstki zapachowe, itd.), inne muszą być również odpowiednie zestawy wrażeń. Związek pomiędzy tymi zestawami może być wyłącznie zewnętrzny, np. asocjacyjny (styczność w czasie i przestrzeni). A zatem, żadne rozróżnienia percepcyjne uzyskane przy pomocy jednych zmysłów nie dadzą się bezpośrednio przełożyć na rozróżnienia wykorzystujące materiał innych zmysłów. Należy dopiero w dalszym przebiegu doświadczenia wytworzyć odpowiednie połączenia pomiędzy ideami.



W *Nowych rozważaniach dotyczących rozumu ludzkiego* Leibniz udziela na pytanie Molineaux przeciwnej odpowiedzi niż Locke: “Założywszy u niewidomego wiedzę, że te dwie figury, które widzi, są sześcianem i kulą, będzie mógł – tak sędzę – rozróżnić je i orzec bez dotykania: to jest kula, a to sześcian. (...) umieściłem [w swojej odpowiedzi] warunek, który uznać można za zawarty w pytaniu; otóż chodzi tylko o samo rozróżnienie i o to, by niewidomy wiedział, że dwie bryły które ma odróżnić, są przed nim, że więc każdy z wygłądów, który widzi, jest bądź wygładem sześcianu bądź wygładem kuli. W tym wypadku wydaje mi się niewątpliwe, że niewidomy, który właśnie przestał nim by, może je rozróżnić przy pomocy zasad rozumowych dodanych do poznania zmysłowego, którego dostarczył mu przedtem dotyk (...) podstawą mego poglądu jest to, że w kuli nie ma żadnych punktów wyróżnionych (...) podczas gdy sześcian ma osiem punktów różnych od wszystkich innych” (1955, t. 1, s. 140-142). Leibniz rozłącza jak widać dwa aspekty percepcji, które w koncepcji Locke'a były połączone: wrażenia zmysłowe i wykrywanie różnic pomiędzy przedmiotami. Dla Locke różnice dane były zmysłowo, dla Leibniza są abstraktami powstającymi na gruncie dowolnej informacji zmysłowej i dlatego dostępne przy pomocy każdej innej informacji zmysłowej, która egzemplifikuje minimalne cechy formalne wymagane przez ten abstrakt (w omawianym przypadku minimum to stanowi posiadanie przez widzianą bryłę wyróżnionych punktów).

Zagadnienie dyskutowane przez Locke'a i Leibniza ma zresztą dłuższą historię. Wprowadził go już Arystoteles w znanym fragmencie rozpoczynającym *Metafizykę*, w którym wiąże percepcję wzrokową z wykrywaniem różnic. “Ludzie z natury dążą do poznania, czego dowodem jest ich umiłowanie zmysłów (...) Nie tylko bowiem gdy działamy, ale nawet wtedy, gdy nie mamy nic praktycznego na względzie, stawiamy wzrok ponad wszelkie inne zmysły. Przyczyna zaś jest to, że ze wszystkich zmysłów wzrok w najwyższym stopniu umożliwia nam poznanie i ujawnia wiele różnic” (1983, s. 3). Nie jest całkiem jasne, czy sformułowanie Arystotelesa sugeruje, że to sam wzrok wykrywa różnice, czy też jedynie stanowi podstawę dla abstrakcyjnego ujęcia. Ta niejednoznaczność kryje właśnie problem, który po dwóch tysiącleciach postawił w sposób wyraźny Molineaux.

Problem pozostaje żywy. Współczesna jego interpretację przedstawił Evans (1985). Jego odpowiedź jest pozytywna, lecz z innego powodu niż odpowiedź Leibniza. Evans twierdzi, że własności przestrzenne nie są przywiązane do określonych danych percepcyjnych (jak chciał Locke), lecz zarazem nie sądzi, by były one abstraktami “wykrywanymi rozumem” (jak chciał Leibniz). Sądzi raczej, że informacja przestrzenna [*spatial content*] dana jest przez zdolności sensomotoryczne. Ktoś, kto był od urodzenia niewidomy nabył, na bazie spostrzeżeń pochodzących z innych zmysłów, pewnej liczby zdolności sensomotorycznych koniecznych do funkcjonowania w przestrzeni. Można powiedzieć, że zdolności te kodują informację o przestrzeni. Inaczej mówiąc własności przestrzenne nie są dane wraz z danymi percepcyjnymi. W czasie poruszania się w przestrzeni wykorzystujemy kompleksy informacji pochodzące z różnych zmysłów. Na ogół sobie tego nie uświadamiamy, lecz można się o tym przekonać bez specjalnych badań, odcinając informację z określonego zmysłu i obserwując deficyt informacji przestrzennej w czasie poruszania się (np. instynktowne częstsze rozglądanie się podczas przechodzenia przez jezdnię zakrytymi uszami). Ponieważ żaden zmysł nie jest tu wyróżniony (wszystkie jedynie przyczyniają się do powstania obrazu przestrzeni, nie ustanawiając samodzielnie żadnej jakości przestrzennej), można powiedzieć, że informacja o przestrzeni jest nadawana informacji zmysłowej przez to, że informacje wejściowe z danego zmysłu mają połączenie z informacjami wyjściowymi dotyczącymi poruszania się ciała w przestrzeni. Wszelkie własności przestrzenne (okrągłość, ciągłość, itd.)



odniesione są do możliwych ruchów ciała. Przestrzeń możliwych ruchów można właśnie nazwać sprawnością [*skill*], o której pisze Evans.

Pogląd Evansa spotkał się z krytyką Millikan (1991). Jeśli Evans podziela odpowiedź Leibniza, twierdząc, że własność informacja nadana informacji zmysłowej może być ponownie rozpoznana, to przyjmuje założenie, że treść przestrzenna jest zakodowana w danej informacji zmysłowej jako pewna własność tej informacji, a zatem jej nośnika, czymkolwiek on jest. Millikan nazywa to pomieszaniem nośnika z treścią. Jeśli bowiem Evans twierdzi, że informacje zmysłowe przyczyniają się wszystkie do powstawania pewnej wspólnej informacji o przestrzeni, to muszą w analogiczny sposób kodować pewną informację przestrzenną. Różnica pomiędzy Evansem a Lockiem, polegałaby zatem tylko na tym, że to nie asocjacje, które muszą dopiero powstać, lecz sprawność, która już jest wykształcona, wyznacza ramy tej wspólnej informacji.

Odpowiedzi na krytykę Millikan próbuje udzielić Grush, wykorzystując do tego kilka eksperymentów myślowych. Na ich użytek wprowadza sygnowane przez specjalną pisownię rozróżnienie na **spostrzeganie** i ***sposprzeganie**. Pierwsze jest posiadaniem reprezentacji, drugie percepcyjnym odnotowywaniem bez budowania reprezentacji. Jedno i drugie zarazem oznacza jako **SPOSTRZEGANIE**. Trochę niefortunne jest to, że na użytek dyskusji o przestrzeni Millikan i Grush używają jako przykładu własności pulsowania, która nie jest w ścisłym sensie własnością trójwymiarowej przestrzeni. Chodzi im o ilustrację problemu na stosunkowo prostym przykładzie. Eksperymenty myślowe są następujące:

Eksperyment 1: Osobnik **A** ma taką konstrukcję układu nerwowego, że kiedy kładzie rękę na przewodzie z płynącym prądem o pulsującym napięciu, odczuwa ciepło proporcjonalne do częstotliwości tego pulsowania prądu. Odczucie ciepła jest wprost proporcjonalne do częstotliwości prądu. Natomiast samo pulsowanie napięcia jest tu tylko ***sposprzegane**. Nie może ono stanowić podstaw do empirycznej generalizacji. Osobnik taki nie mógłby na przykład wykonać na przykład ruchów dyrygenta do słyszanego rytmu. ***Sposprzeżenie** rytmu pulsującego prądu nie przekłada się na określoną „kompetencję” ciała ani na żadną kompetencję poznawczą.

Eksperyment 2: Osobnik **B** nosi od urodzenia bez przerwy okulary, które prezentują mu różne bodźce wizualne, lecz nie są one powiązane z własnościami jego otoczenia i działania. Inaczej mówiąc, osobnik ten nie nabiera żadnych sprawności na podstawie informacji wizualnej, jaka jest mu dostarczana. Grush sugeruje, że osobnik ten ma wprawdzie spostrzeżenia wzrokowe, lecz nie zawierają one informacji o przestrzeni. Grush sugeruje, (wbrew, na przykład, dawnym fenomenologicznym ustaleniom Husserla), że stosunek pomiędzy wrażeniem wzrokowym a pewną porcją przestrzeni jest przygodny, nie zaś konieczny. Związek ma tu być analogiczny do związku pomiędzy percepcją ojczystego języka a odbieraniem dźwięków jako sensownych. Związek jest natychmiastowy, lecz nie konieczny, tylko zdeterminowany moim wyuczeniem się języka. To, iż ktoś może mieć wrażenia wzrokowe nie zawierające informacji przestrzennej wydaje się dziwacznym wnioskiem. Broniąc jego poprawności Grush analizuje trzeci eksperyment myślowy.

Eksperyment 3: Znany jest wynalazek dla niewidomych, sonar, który emituje dźwięki niesłyszalne i odbiera echo odbite od przedmiotów a następnie przetwarza to echo w dźwięki słyszalne, tak by różne dźwięki odpowiadały różnym odczytom echa. Urządzenie okazuje się dość skuteczne: informacje zakodowane jako dźwięki odtwarzają wysoce wiarygodną przestrzeń – podobną do tej, jaką widzi wzrok. Oddaje nie tylko odległości i wielkości



przedmiotów, lecz nawet tak subtelne własności przestrzenne, jak faktura. Możemy sobie wyobrazić osobnika **C**, który od urodzenia używa bez przerwy takiego sonaru. Można powiedzieć, że dźwięki niosłyby dla takiego człowieka informację przestrzenną. Osobnik ten mógłby się dziwić, że dla pozostałych ludzi jest inaczej (dla nas pojedyncze dźwięki nie mają w sobie nic przestrzennego, choć pewnego rodzaju przestrzeń występuje przy skomplikowanych układach dźwiękowych) – sam nie mógłby bowiem nawet, gdyby chciał, oderwać wrażenia słuchowego od informacji przestrzennej, np. “duży przedmiot po lewej stronie”, “Okolo 2 metrów do ściany”, itd. Różnica pomiędzy osobnikiem **B** a osobnikiem **C** polega na tym, że pierwszy ma wrażenia wzrokowe bez treści przestrzennej (Grush sugeruje nawet, że nie mógłby odróżnić dwóch identycznych a tylko inaczej umieszczonych w przestrzeni plam barwnych).

Na gruncie tych rozważań Grush wprowadza rozróżnienie na informację o przestrzeni, która może być – jak chciał Locke – kodowana w poszczególnych percepcjach zmysłowych, i **treść przestrzenną** [*spatial content*], która zależy od tego, czy na gruncie danej informacji rozwinęły się określone sprawności psychomotoryczne.

Przedstawiona powyżej dyskusja ma charakter czysto analityczny. Jej przedmiotem jest ustalenie, czy istnieją powody do przyjęcia i posługiwania się pojęciem międzymodalnej informacji przestrzennej. Same jednak argumenty spekulatywne nie są rozstrzygające, wskazują tylko na problem i konieczność pewnych pojęciowych rozróżnień. Same pojęcia nie stanowią jednak teorii, wyznaczają tylko jej elementy składowe. W kolejnym paragrafie zastanowię się nad tym, czy są teoretyczne powody do uznania, że głębia widzenia konstituowana jest przez elementy niewizualne, a nawet w ogóle niepercepcyjne, jak na przykład sprawności psychomotoryczne. W tym celu wykorzystam klasyczną już dziś teorię widzenia D. Marra.

2. Teoria widzenia D. Marra i jej ograniczenia

Przejście od szkicu pierwotnego do szkicu 2 1/2 opiera się na tzw. efekcie stereoskopowym, czyli mierzeniu przez aparat wzrokowy rozbieżności pomiędzy obrazami na obu siatkówkach. Operacja ta wymaga stosunkowo skomplikowanego mechanizmu przeliczającego. Ma on dwie główne cechy: **(1)** algorytmy są kooperatywne tzn. dają pożądany rezultat jedynie na skutek zaangażowania wielu niezależnych algorytmów. **(2)** algorytmy składowe działają lokalnie (realizują je określone grupy komórek nerwowych), dlatego całościowa organizacja obrazu dwu i półwymiarowego musi przebiegać w innym miejscu, na poziomie obrazu siatkówki reproduktowanego w korze wzrokowej (wyspecjalizowane komórki pola VI).

W swojej teorii widzenia Marr (1982) proponuje hierarchiczny model powstawania przetwarzania informacji wizualnej, począwszy od rozkładu i typu pobudzenia komórek siatkówki do reprezentacji kształtów i odległości w trójwymiarowej przestrzeni. Podstawowym mechanizmem tego przetwarzania są funkcje działających jak filtry częstotliwości (zob. Marr 1982, s. 54-61). Funkcje te są implementowane w algorytmach realizowanych dzięki określonym własnościom komórek nerwowych. Ustalając regiony o różnej jasności funkcje te pozwalają na wykrywanie krawędzi oddzielających te regiony, co stanowi podstawowy krok układu wzrokowego w odtwarzaniu przestrzennych własności otoczenia. Ten produkt przekształcenia (pole jasności z krawędziami) nazywa Marr szkicem pierwotnym. Z niego aparat wzrokowy otrzymuje stereoskopowy szkic, wykorzystujący obie



siatkówki – Marr nazywa go szkicem 21/2. Zawiera on informację o głębi, lecz ujętą niejako w perspektywicznym skrócie: trzeci wymiar nie jest tu w pełni rozwinięty. Obraz jest “zrośnięty” z punktem widzenia i dlatego nie można nim swobodnie manipulować, jako wewnętrzną reprezentacją.

Podstawowy problem polega na tym, że zmierzenie rozsunęcia wymaga ustalenia, które punkty odpowiadają sobie na obu siatkówkach. Teoretycznie bowiem aparat wzrokowy mógłby zestawiać informacje (czyli oceniać je pod kątem zbieżności lub niezbieżności obrazów) pochodzące z wielu różnych miejsc na obu siatkówkach. Wiązki linii widzenia biegnące od jednej i drugiej siatkówki do danego obiektu w odległości D przecinają się bowiem w wielu miejscach w obszarze wyznaczonym przez $D+a$ i $D-b$ (a i b zależą od wielkości obiektu), jednak te punkty przecięcia niosą w większości fałszywe informacje o położeniu widzianego obiektu w rzeczywistej przestrzeni. Tylko jeden zbiór przecięć niesie prawidłową informację i tylko on powinien być podstawą mierzenia rozbieżności.

Żadne bezpośrednie połączenie neuronalne pomiędzy komórkami jednej i drugiej siatkówki nie może zagwarantować ustalenia właściwej korespondencji pomiędzy obrazami na siatkówkach; system widzenia stereoskopowego nie może mieć reprezentacji obrazów nierozbieżnych, ponieważ kiedy mamy otwarte oboje oczu nasza kora wzrokowa odnotowuje obrazy rozbieżne. Oczywiście możliwe jest, że informacja o rozsunieciu nakłada się na pewien utrwalony model korespondencji pomiędzy grupami komórek obu siatkówek i zapewne tak właśnie jest. Nie rozwiązuje to jednak problemu “ślepych rozwiązań” wzmiankowanych przez Marra, czyli takich przecięć linii widzenia biegnących z jednego i drugiego oka, którym nie odpowiadają żadne przedmioty w rzeczywistej przestrzeni. Z drugiej strony, powtórzmy, pewna ustalona korespondencja pomiędzy obrazami na obu siatkówkach jest konieczna jako podstawa do mierzenia rozbieżności tych obrazów. Wydaje się więc, że aparat wzrokowy musi dysponować czymś w rodzaju wirtualnej odpowiedniości obrazów jednoocznych, wobec której mierzy przesunięcie.

Wirtualna odpowiedniość kodowanych w obu siatkówkach jest czymś abstrakcyjnym. W celu skonstruowania i zachowywania w czasie percepcji tego abstrakcyjnego parametru system widzenia głębi musiałby go niejako stale poszukiwać. Mogłoby to zachodzić przez ustawiczne uzmiennianie obrazów na siatkówce. Na skutek uzmienniania rozkład pobudzeń siatkówki miałby tendencję do oscylowania wokół pewnej wartości wypadkowej. Cały proces odbywałby się już na bardzo rudymentalnym poziomie, przy pomocy automatycznych ruchów gałki ocznej, a także ruchów głowy i całego ciała, oraz przez zmiany w ogniskowaniu wzroku.

Ruch gałek ocznych jest istotny również dlatego, że mierzenie rozbieżności jest najskuteczniejsze wtedy, gdy wstępna obróbka informacji na poziomie dwuwymiarowego obrazu (widzenie jednooczne) odbywa się przy pomocy kanałów o największej rozdzielczości. Fizjologicznie kanały te nie są dostępne w dowolnym punkcie siatkówki, ale tylko w niektórych miejscach i co za tym idzie, muszą być czynnie skierowane na pole wizualne. Aparat wzrokowy celowo przemieszcza i włącza do pracy owe kanały. Marr pisze: “[System widzenia dwuocznego] opiera się na ruchu gałek ocznych w celu zbudowania całościowej i dokładnej mapy niewspółmierności z obu punktów widzenia. Ruchy te są niezbędne, ponieważ najbardziej dokładne pomiary niewspółmierności osiągane są z kanałów o największej rozdzielczości; ruch oka sprawia, że każda część sceny w końcu znajduje się w zakresie widzenia, od których wychodzą owe kanały o największej rozdzielczości” (Marr 1982, s. 128).



Z powyższych sugestii Marra (uzupełnionych przeze mnie sugestią o potrzebie ustalania korespondencji pomiędzy obrazami jednoocznymi **przed** zmierzeniem rozbieżności), wynika, że potrzebne są dodatkowe mechanizmy wspierające stereoskopię (mechanizm widzenia głębi oparty na mierzeniu rozbieżności pomiędzy obrazami na siatkówkach). I rzeczywiście, wiadomo dziś z dużą pewnością, że “poza stereoskopią istnieją jeszcze inne mechanizmy wykrywania głębi. Większość z nich służy do oceny stosunków przestrzennych między obiektami leżącymi w dalszej odległości od obserwatora. Do najważniejszych należy wykorzystywanie informacji o konwergencji (kącie patrzenia) oczu, o deformacjach perspektywicznych (linie równoległe wydają się zbieżne), o nakładaniu się obrazów (obrazy bliższe przysłaniają dalsze) oraz o wielości obrazów na siatkówce (obraz obiektów dalszych jest mniejszy). Są również mechanizmy opierające się na informacjach o ruchu” (Górska i inni, 2000, s. 167).

Najważniejszy wszakże wydaje mi się pewien mechanizm globalny. Patrzący przedmiot musi być w ruchu; zmienia się położenie jego ciała nawet, jeśli wydaje się prawie nieruchome (lekkie zmiany punktu widzenia mogą być wystarczające dla wykonania przez aparat wzrokowy odpowiedniej ilości wirtualnych przekształceń; również gałki oczne przez cały czas wykonują dość skomplikowane ruchy. Informacja o trójwymiarowej przestrzeni osiągana przez penetrację dzięki ruchom gałki ocznej i całego ciała stanowi integralną część procesu przetwarzania obrazu na siatkówce, prowadzącego do stereoskopii. Wizualna głębia jest więc wprawdzie wywiedziona z przekształcenia płaskiego obrazu na siatkówce, ale raczej samo to przekształcenie jest oparte na pierwotnie danej trójwymiarowej przestrzeni czuciowo-ruchowej.

Wiele wskazuje na to, że proces ustalania odpowiedniości a zarazem rozbieżności pomiędzy obrazami na obu siatkówkach musi korzystać z niezależnej informacji o rzeczywistości. Ta informacja nie jest oczywiście jakąś inną informacją wizualną - nie istnieje przecież żadne super-widzenie, które sterowałoby “normalnym” widzeniem. Istnieje jednak model wizualnej przestrzeni, który decyduje o granicach, tempie, rytmie uzmienniania pola wizualnego i w ten sposób jest źródłem parametrów sterujących widzeniem. Skąd bierze się taki model? Wydaje mi się sensowną sugestią – niestety znacznie to mniej niż naukowa hipoteza – że ruchy gałek ocznych i całego ciała – nieustanna, choć często nieuświadomiana aktywność patrzącego podmiotu, generuje serię próbnych przestrzeni wizualnych, które następnie ulegają swoistej syntezie. Sugeruję (również w Piłat, 1999), że system kontroli aktywności ciała jest logiczny i funkcjonalnie pierwotny w stosunku do informacji wizualnej. Widzenie głębi czerpie informacje nie tylko z oświetlonych powierzchni przedmiotów, którym odpowiadają odpowiednie siły pobudzeń komórek siatkówki, ale również z wirtualnych przestrzeni (modeli przestrzennych) formowanych przez eksplorujące ruchy gałek ocznych sprzężony z pozycją i ruchem ciała.

3. Fenomenologia głębi

Opis widzenia trójwymiarowego przedstawiony przez Merleau-Ponty’ego najpierw w *Fenomenologii percepcji* (1945) a później w eseju *Słowo i obraz* (1996), pokazuje, że percepcja głębi jest nieodłączna od ruchu ciała w rzeczywistej trójwymiarowej przestrzeni. Pomiedzy ruchem ciała a elementami wizualnej przestrzeni istnieje stała interakcja. Nie można tu mówić o elementach pierwotnych i wtórnych. Ruch ciała zmienia wizualną przestrzeń, a zmiana tej ostatniej zmienia sytuację ciała i jego ruch. Ta interakcja sprawia, że



przestrzeń, która prezentuje się naszym oczom jest tą samą przestrzenią, które zajmuje nasze ciało. Wzajemna relacja całego ludzkiego ciała i przestrzeni dochodzi do głosu w malarstwie, które, jak stara się pokazać Merleau-Ponty, nie jest prostym rzutowaniem widzianej przestrzeni na płaszczyznę, ale oddaniem interakcji pomiędzy ciałem a przestrzenią wizualną. Czytamy: “[Głębia] ma w sobie coś paradoksalnego: oto widzę przedmioty, które są zakryte przez przedmioty inne, których wcale nie widzę, skoro są ukryte jedne za drugimi; ja też widzę, jakkolwiek jest niewidoczna, liczy się bowiem od mojego ciała, by skończyć na rzeczach, my zaś jesteśmy z nią stopieni (...) gdy patrzę na przedmioty z boku, wydaje mi się, że je widzę w rozstawieniu, to dlatego, że nie zasłaniają się całkiem - widzę je jedne spoza drugich według inaczej liczonej szerokości. Zawsze jest się z tej lub tamtej strony głębi. Nigdy rzeczy nie są jedne za drugimi. Wyjście z ukrycia i pozostawanie w nim nie wchodzi do definicji rzeczy, wyrażają tylko moją niepojętą współzależność, zupełnie niezrozumiałą solidarność z jedną z nich, mianowicie z moim ciałem, i wszystko, co jest pozytywnego w tym ukryciu i w wyjściu z niego, należy do urabianych przeze mnie myśli, a nie do atrybutów rzeczy” (Merleau-Ponty 1996, s. 38-39).

Relacje konstytuujące głębię rodzą się ze związku ciała z przestrzenią i dlatego są czymś pierwotnym, a nie czymś konstruowanym. “Tajemniczość głębi, powiada Merleau-Ponty, nie jest pozbawionym tajemnicy odstępem, który dostrzegłbym pomiędzy drzewami z samolotu; ani też skrywaniem się jednych rzeczy za drugimi, co oddaje rysunek perspektywiczny. Oba te widoki są aż nadto wyraźne i nie stanowią jakiegokolwiek problemu. Zagadką jest ich powiązanie, czyli to, co jest między nimi (mianowicie to, że spośród rzeczy każdą widzę na swoim miejscu właśnie dlatego, że jedna zasłania drugą - to, że rywalizują ze sobą przed moim spojrzeniem, ponieważ są na swoim miejscu). (...) O tak pojmowanej głębi nie można już mówić, że jest ‘trzecim wymiarem’. Przede wszystkim, gdyby była wymiarem, byłaby raczej pierwszym” (Merleau-Ponty 1996, s. 50).

Opis Merleau-Ponty'ego uzmysławia, że w każdej sytuacji percepcyjnej ma się dwa punkty widzenia jednocześnie zaangażowane w widzenie głębi. Pierwszy powstaje przez "solidarność z własnym ciałem", dzięki niemu jest się od razu w trójwymiarowej przestrzeni, choć przestrzeń ta jest egocentryczna. Drugi punkt widzenia jest zewnętrzny wobec tego, co widziane; pozwala zerwać z egocentrycznością, lecz sprawia, że trójwymiarowość staje się problematyczna; musi być dopiero zrekonstruowana przy pomocy różnych mechanizmów, w jakie wyposażony jest ludzki aparat wzrokowy. Oba punkty widzenia są jednak zakorzenione w ciele, a dokładnie w percepcji własnego ciała. Swoje ciało widzimy zawsze częściowo. To, że w ogóle je widzimy, daje nam poczucie przebywania "z tej strony głębi", pierwotną informację przestrzenną. Wszakże to, że widzimy je tylko częściowo sprawia, że nasz percepcyjny punkt widzenia, nie jest klarowny – patrzenie staje się podobne do myślenia. Tak właśnie - spekulatywnie - rzecz ujmował Kartezjusz. Merleau-Ponty oddaje jego pogląd następująco. "Gdy dla przykładu chce się zrozumieć, jak dochodzi do tego, że widzimy położenie przedmiotów, nie ma innego sposobu niż założenie takiej duszy, która wiedząc, gdzie są konkretnie części jej ciała, zdolna jest "swoją uwagę przenieść stąd" na wszystkie punkty przestrzeni, znajdujące się na przedłużeniu poszczególnych członków (Kartezjusz, *Dioptrique*, VI). Następnie Merleau-Ponty komentuje: "Lecz jest to tylko jeden model zdarzenia. W jaki bowiem sposób zdobywa ona wiedzę o tej przestrzeni własnego ciała, którą rozciąga na rzeczy, o tym pierwszym tutaj, skąd wszystkie tam pochodzą będą?" (M. Merleau-Ponty, *Oko i umysł*, s. 43).

Filozoficzna i psychologiczna wykładnia widzenia głębi była jak dotąd, zdaniem Merleau-Ponty'ego, zdominowana przez odkrycie perspektywy i związaną z nim geometryzację



przeżywanej przestrzeni (stąd geometryzacja, a zatem intelektualizacja ludzkiego ciała). Wynalazek malarski zbiegł się z odkryciem Kartezjusza tworząc pewien stereotyp przestrzeni wizualnej. Merleau-Ponty ujmuje to następująco: “Dla duszy ciało jest jej rodzinną przestrzenią oraz macierzą pozostałej macierzy istniejącej. W ten sposób widzenie podwaja się - jest takie widzenie, nad którym się namyślam, które mogę tylko jako myślenie pomyśleć, jako inspekcję umysłu, osąd, czytanie znaków. I jest też widzenie, które ma właśnie teraz miejsce którego ideę można zdobyć tylko wtedy, gdy się praktykuje samemu i które wprowadza pomiędzy przestrzeń i myślenie autonomiczny porządek właściwy złożeniu duszy i ciała. Zagadka widzenia nie została wyeliminowana: została przeniesiona z widzącego myślenia do aktu widzenia” (M. Merleau-Ponty 1996, s. 43).

[Kartezjusz] miał słuszość, gdy inspirował się perspektywicznymi technikami Renesansu, albowiem zachęciły one malarstwo do swobodnego przeprowadzania doświadczeń z głębią, a w ogólności dokonywania najrozmaitszych prezentacji Bytu. Były one fałszywe jedynie w tej mierze, w jakiej zamierzały zamknąć sobą okres poszukiwań i dzieje malarstwa, doprowadzić do dokładnego i bezbłędnego malarstwa... Spośród technik perspektywicznych żadna nie jest rozwiązaniem dokładnym, nie ma rzutów oddających wiernie pod każdym względem świat istniejący i godnych stania się podstawowym prawidłem czy prawem malarstwa... Rzuty płaskie... Przekroczywszy pewien stopień deformacji, odsyłają na powrót do naszego punktu widzenia, a co się tyczy samych rzeczy, giną w oddaleniu, którego nie pokona żadne myślenie (M. Merleau-Ponty 1996, s. 41-42).

Tak samo, jak malarstwo przyczyniło się do wyidealizowania przestrzeni wizualnej, tak też ono samo walczy o odzyskanie prawdziwej fenomenologii głębi. Merleau-Ponty odwołuje się szczególnie do sztuki Paula Cezanne'a. “Ja myślę, że Cezanne przez całe życie poszukiwał głębi (...) Kiedy Cezanne szuka głębi, to szuka deflacji Bytu, ona zaś występuje we wszystkich trybach przestrzeni, jak również w formie. Już Cezanne wie o tym, co powtórzy kubizm: że forma zewnętrzna otoczka, jest wtórna, pochodna, że nie jest tym, co sprawia, iż dna rzecz nabiera kształtu, że trzeba rozbić przestrzenną skorupę, przełamać kompotierkę i w to miejsce malować... Co? Sześciany, kule, stożki, jak to raz powiedział? Czyste formy posiadające stałość przynależną czemuś, co określone może być wewnętrznym prawem konstrukcji... A więc przestrzeni i jej zawartości trzeba szukać jednocześnie. Problem się uogólnia. Nie jest to już problem odległości i linii oraz formy, lecz jest też problem koloru. Jest on miejscem, gdzie nasz mózg łączy się ze wszechświatem” (*Oko i umysł*, s. 51-52). Przestrzeń przestaje być trójwymiarowym pojemnikiem na rzeczy, staje się kompozycją jakości i relacji pomiędzy rzeczami.

Tak samo, jak perspektywa, również tradycyjne traktowanie konturu przedmiotów musiało ulec pod presją odkryć nowego malarstwa. “Kontur przedmiotów, pojmowany jako okalająca je linia, nie jest zjawiskiem ze świata widzialnego, lecz z geometrii. Jeśli zakreślimy linią kontur jabłka, zrobimy zeń rzecz, podczas gdy w istocie kontur jest granicą idealną, ku której biegną w głąb brzegi jabłka. Gdybyśmy w ogóle nie zaznaczyli konturu, odebralibyśmy przedmiotom tożsamość. Zaznaczając tylko jeden, musielibyśmy zrezygnować z głębi, tj. z wymiaru wyznaczającego rzecz, która nie odkrywa się przed nami całkowicie, lecz pełna potencjalnych możliwości pozostaje rzeczywistością niewyczerpaną. Dlatego też Cezanne modulowaniem barw oddawać będzie krągłość przedmiotu i niebieską linią zaznaczy wiele konturów. Spojrzenie biegnące od jednego do drugiego uchwyci kontur tworzący się z nich wszystkich, tak, jak to się dzieje w postrzeganiu. Nie ma nic arbitralnego w tych słynnych deformacjach...” (M. Merleau-Ponty 1996, s. 80).



Poszukiwania Cezanne'a w dziedzinie perspektywy doprowadziły, dzięki wiernemu przedstawieniu zjawisk, do odkrycia tego, co sformułowała psychologia współczesna. Perspektywa żywa, perspektywa naszego spostrzegania nie jest perspektywą geometryczną czy fotograficzną. W percepcji przedmioty bliższe wydają się mniejsze, a przedmioty oddalone większe niż na fotografii, i tak na przykład w kinie nadjeżdżający pociąg rośnie w oczach znacznie szybciej, niż to się dzieje w rzeczywistości. (s. 78)

Przełamanie w nowoczesnym malarstwie tradycyjnej perspektywy i konturu wiąże Merleau-Ponty z rolą ciała w widzeniu głębi, o którym była mowa wcześniej w tym paragrafie. "Malarz 'wnosi swoje ciało', powiada Valery. I rzeczywiście trudno sobie wyobrazić, żeby Umysł mógł malować. Oto użyczając swego ciała światu, malarz przemienia świat w malarstwo. Aby zrozumieć te transsubstancjacje, trzeba odnaleźć aktualne i operatywne ciało, nie to, które jest wycinkiem przestrzeni, pękiem funkcji, ale które jest splotem widzenia i ruchu. (1996, s. 20).

Powiedzenie "malarz wnosi swoje ciało" zostało z największą dosłownością potraktowane przez współczesnego artystę amerykańskiego Dawida Hockney'a, który w związku ze swoimi eksperymentami fotograficznymi zauważył, że zdjęcie nie zawiera owej charakterystycznej dla obrazu obecności cielesnej aktywności malarza widocznej w śladach pracy pędzla, czy innego narzędzia malarskiego. Pragnąc wprowadzić ten element do fotografii składał swe prace z wycinków różnych zdjęć naklejanych obok siebie i częściowo zachodząc na siebie. Oprócz interesujących rezultatów czysto graficznych, Hockney osiąga ten efekt, że praca rąk twórcy staje się na jego fotografiach widoczna – fotografia przestaje być jedynie wytworem pozaludzkiego systemu optycznego, do którego człowiek wnosi jedynie kadrowanie.

4. Głębia widzenia i model świata

Problemy, na jakie napotykamy chcąc wyjaśnić strukturę fenomenalną widzenia głębi i odpowiedzialne za nią mechanizmy biologiczne pochodzą stąd, że tylko z pozoru chodzi tu o coś prostego: szacowanie odległości przedmiotów od własnego ciała i wzajemnie jednych od drugich. W istocie chodzi o niezmiernie skomplikowany system kontroli ruchu w przestrzeni i symulowanie tego ruchu w postaci pewnej wewnętrznej reprezentacji. Ponadto wśród własności przestrzennych wykrywanych wzrokiem są nie tylko odległości, lecz także wymiary, kształty i faktury powierzchni. Ważne miejsce we wzrokowym postrzeganiu przestrzeni ma również przestrzenne umiejscowienie własnego ciała, które jest czymś pierwotniejszym od oceny odległości przedmiotów – trzeba je wszakże mierzyć do jakiegoś punktu, tymczasem nie sposób wskazać konkretnego punktu fizycznego w ciele człowieka, który służy mu (podobnie jak słup umieszczony w centrum Warszawy, na którym umieszczono napisy informujące o odległości od centrów innych miast) jako punkt odniesienia w szacowaniu odległości dzielącej ciało tego człowieka od przedmiotów.

Widzenie głębi zakłada funkcjonowanie wielu mechanizmów o różnym stopniu automatyzacji. Niektóre z nich – jak ogniskowanie wzroku pozwalające szacować odległość – są po prostu wbudowane w działanie oka. Inne jednak dotyczą stosunkowo wysokiego poziomu przetwarzania informacji. Przez wysoki poziom przetwarzania rozumiem ustalanie informacji przy pomocy innych informacji, które na najbardziej elementarnym poziomie są system różnic stanów komórek i grup komórek układu nerwowego odpowiedzialnego za widzenie. Odnotowana przez system – czyli relewantna – różnica jest dopiero informacją.



Różnice pomiędzy różnicami stanowią informacje wyższego (i odpowiednio coraz wyższych rzędów). Wiele wskazuje na to, że w ustalaniu informacji wzrokowych wyższych rzędów bierze udział wiele mechanizmów sterujących i [poznawczych nie należących w ścisłym sensie do układu wzrokowego. Są to między innymi: schemat własnego ciała oraz możliwość budowania wirtualnych przestrzeni w celu przeprowadzenia symulacji ruchu własnego i ruchu spostrzeganych przedmiotów. To jak zbudowane są te wirtualne przestrzenie, a także to jak “wygląda” u danej osoby model jej własnego ciała zależy od jej **osobistego modelu świata** tej osoby. To, co Merleau-Ponty nazwał za Valery’em “wnoszeniem własnego ciała” dotyczy każdego widzącego człowieka i jest swoiste dla każdego człowieka. Mamy zarówno wrodzone jak nabyte skłonności do organizowania sobie wizualnej przestrzeni, która może być mniej lub bardziej egocentryczna, mniej lub bardziej zlateralizowana (preferencja jednej ze stron), mniej lub bardziej ciągła, o większej lub mniejszej głębokości” czasowej, czyli zdolności do antycypującego ujęcia.

Przedstawiając składowe funkcjonalne – a na koniec sugerując również pojęciowe – percepcji głębi opowiadam się również za argumentacją Evansa i Grusha (niezależnie od niewielkiej siły przekonywania eksperymentów myślowych tego ostatniego) za tezą o ścisłym związku pomiędzy sprawnością motoryczną a wizualną informacją o przestrzeni.

Wrzesień 2001

Literatura:

- [1] Arystoteles, *Metafizyka*, tłum. K. Leśniak, Warszawa 1983.
- [2] R. Grush, *Skill and Spatial Content*, w: “Electronic Journal of Analytic Philosophy”, <http://www.phil.indiana.edu/ejap/1998/grusharticle98.html>, 1998.
- [3] G.W.F. Leibniz, *Nowe rozważania dotyczące rozumu ludzkiego*, tłum. I. Dąbska, t 1-2, Warszawa 1955.
- [4] J. Locke, *Rozważania dotyczące rozumu ludzkiego*, tłum. B. Gawecki, Warszawa 1955.
- [5] M. Maciejczak, *Świat według ciała. Fenomenologia percepcji M. Merleau-Ponty’ego*, Toruń 1995.
- [6] D. Marr, *Vision*, San Francisco 1982.
- [7] M. Merleau-Ponty, *Oko i umysł. Szkice o malarstwie*, oprac. S. Cichowicz, Gdańsk 1996.
- [8] *Mózg i zachowanie*, red. T. Górską, A. Grabowska, J. Zagrodzka, Warszawa 2000.
- [9] R. Piłat, *Umysł jako model świata*, Warszawa 1999.