

» Idź do

- Spis treści
- Przykładowy rozdział

» Katalog książek

- Katalog online
- Zamów drukowany katalog

» Twój koszyk

- Dodaj do koszyka

» Cennik i informacje

- Zamów informacje o nowościach
- Zamów cennik

» Czytelnia

- Fragmenty książek online

» Kontakt

Helion SA
ul. Kościuszki 1c
44-100 Gliwice
tel. 032 230 98 63
e-mail: helion@helion.pl
© Helion 1991-2008

Kompresja dźwięku i obrazu wideo (Real World)

Autor: Andy Beach
Tłumaczenie: Jacek Janusz
ISBN: 978-83-246-1898-9
Tytuł oryginału: [Real World Video Compression \(Real World\)](#)
Format: 172x245, stron: 352



Zapewnij najlepszą jakość przy najmniejszym rozmiarze!

- Jak dobrać optymalne parametry?
- Jakich narzędzi potrzebujesz?
- Jak sprawnie skompresować materiał wideo na stronę WWW?

Czy chcesz opublikować film z wakacji na Twojej stronie WWW? Zaraz, zaraz... czy aby na pewno Twoi przyjaciele chcą ściągać tyle danych? Czy masz wystarczającą ilość miejsca na Twoim serwerze WWW? Prawdopodobnie odpowiedź brzmi „nie” – dlatego przed opublikowaniem Twój film należy poddać kompresji. Łatwiej powiedzieć, niż zrobić! Jak ustalić odpowiednie parametry kompresji? Co zrobić, żeby wynikowy plik przy rozsądnym rozmiarze zapewniał najwyższą jakość? Ta książka odpowie na te i wiele innych pytań!

Dzięki książce „Kompresja dźwięku i obrazu wideo” zdobędziesz wiedzę na temat technologii audio i wideo oraz obowiązujących pojęć. Na początku poznasz teoretyczne podstawy kompresji dźwięku i obrazu, pozwalające Ci lepiej przyswoić sobie wiedzę z kolejnych rozdziałów. Po przeczytaniu tego podręcznika będziesz dysponować wiedzą na temat używanych kodeków czy dostępnych parametrów kompresji. Dowiesz się, w jaki sposób skompresować film dla DVD, internetu oraz urządzeń przenośnych. Dodatkowo zorientujesz się wśród dostępnych narzędzi, służących do kompresji plików wideo, takich jak Adobe Media Encoder, QuickTime Pro, Compressor czy też Microsoft Expression Encoder. Jeżeli publikujesz materiały wideo w sieci, koniecznie zapoznaj się z tą książką!

- Podstawy technologii wideo i audio
- Dostępne kodeki
- Ustalanie parametrów kompresji
- Sprzęt zalecany do kompresji
- Przetwarzanie wstępne materiału wideo
- Dostępne narzędzia kompresji
- Kompresja dla DVD
- Zasady kompresji dla internetu
- Kompresja dla urządzeń przenośnych

**Sprawnie kompresuj swoje materiały wideo i publikuj je bez przeszkód,
gdzie tylko zechcesz!**



Spis treści

Wstęp..... 13

O czym ta książka nie jest	14
W takim razie, czego dotyczy ta książka?	14
Kto powinien przeczytać tę książkę?	15
Strona internetowa książki	15
O autorze	16
Podziękowania	16

Rozdział 1. Zrozumieć technologię wideo i audio 19

Elementy wideo	20
Klatki i pola	21
Korzyści ze stosowania przeplotu	22
Problemy podczas stosowania przeplotu	23
Wideo ze skanowaniem progresywnym	25
Rozdzielczości	26
Współczynnik kształtu	28
Sygnały analogowe a cyfrowe	31
W jaki sposób działa kompresja	32
Kompresja stratna i bezstratna	34
Kompresja przestrzenna (DCT) i kompresja wavelet	35
Kwantyzacja	36
Kompresja międzyklatkowa i wewnątrzklatkowa	36
Kompresja dźwięku	37

Szacowanie wideo dla celów kompresji	38
W jaki sposób wideo zostało zapisane?	39
W jaki sposób przeprowadzono edycję wideo?	39
Kto ma być docelowym odbiorcą?	40
Zakończenie	40

Rozdział 2. Język kompresji41

Odtwarzacze	42
Odtwarzacz QuickTime Player	43
Odtwarzacz Windows Media Player	47
Odtwarzacz Adobe Flash Player	49
Odtwarzacz RealPlayer	51
Inne odtwarzacze wideo	53
Kontenery	55
Kodeki	56
Opisy powszechnie używanych kodeków	57
Na wszystko jest czas i miejsce	61
Kodeki finalne	62
Kodeki edycyjne	62
Kodeki przestarzałe	63
Parametry kompresji	65
Prędkości transmisji	66
VBR i CBR	66
Szybkość klatek	67
Rodzaje klatek	68
Współczynniki kształtu	70
Piksele o różnych kształtach	72
Rozdzielczość	72
Udostępnianie przez sieć: strumieniowanie i pobieranie	73
Zakończenie	76

Rozdział 3. Recepty na sukces77

Jakie są Twoje wymagania kompresji?	78
Prezentacje krótkotrwałe czy długotrwałe?	78
Jakie jest Twoje docelowe medium prezentacyjne?	78
W jaki sposób zawartość wideo została dostarczona do Ciebie?	78
Jak wiele treści wideo kompresujesz?	79
Czy kompresja jest jedyną czynnością, którą wykonujesz?	80
Jakość czy dostarczenie we właściwym czasie?	80
Czy jesteś właścicielem zawartości, czy tylko „trybikiem w maszynie”?	80
Jak dużą wiedzę ma Twój klient?	80
Sprzęt i sposób działania dla różnych przypadków	81
Kodowanie jako montażysta	81
Kodowanie jako specjalista kompresji	84

Automatyzacja procesu kodującego	86
Systemy przemysłowe: pierwsza liga	86
Podstawowy sprzęt kodujący	87
Sprzęt	87
Oprogramowanie	91
Wskazówki projektowe	96
Rozsądna minimalizacja utraty jakości	96
Bądź uporządkowany	96
Konwencje nazewnictwa plików	97
Eksperymentowanie może być dobrym pomysłem	98
Stwórz własny zbiór przepisów	99
Pliki pośrednie	99
Archiwizacja i transkodowanie	100
Zakończenie	100

Rozdział 4. Przetwarzanie wstępne 101

Usuwanie przeplotu wideo	102
Łączenie pól (blend)	103
Weave	103
Usuwanie oparte na obszarach (area-based)	103
Rozmywanie obrazu (motion blur)	103
Odrzucanie (discard)	103
Bob	104
Skanowanie progresywne	104
Kompensacja ruchu (motion compensation)	104
Konwersja klatek i odwrócona konwersja klatek	105
Kadrowanie	106
Skalowanie	107
Korekcja współczynnika kształtu obrazu	109
Korekcja współczynnika kształtu piksela	110
Redukcja zakłóceń	111
Regulacja luminancji	112
Korekcja gamma	113
Jasność i kontrast	114
Modyfikacja chrominancji	114
Nasycenie	114
Odcień barwy	114
Wstępne przetwarzanie ścieżki dźwiękowej	115
Regulacja siły dźwięku	115
Redukcja szumów	116
Zakończenie	116

Wywiad z kompresjonistą: John Howell 117

Rozdział 5. Narzędzia kompresji 121

Adobe Media Encoder	122
Jak wygląda ten program?	123
W jaki sposób działa ten program?	125
Czym wyróżnia się ten program?	128
Na co należy zwrócić uwagę podczas używania tego programu?	128
Co powinienś zapamiętać?	129
QuickTime Pro	129
Jak wygląda ten program?	131
W jaki sposób działa ten program?	133
Czym wyróżnia się ten program?	138
Na co należy zwrócić uwagę podczas używania tego programu?	139
Co powinienś zapamiętać?	139
Compressor	140
Jak wygląda ten program?	141
W jaki sposób działa ten program?	143
Czym wyróżnia się ten program?	147
Na co należy zwrócić uwagę podczas używania tego programu?	148
Co powinienś zapamiętać?	149
Microsoft Expression Encoder	149
Jak wygląda ten program?	150
W jaki sposób działa ten program?	152
Czym wyróżnia się ten program?	156
Na co należy zwrócić uwagę podczas używania tego programu?	156
Co powinienś zapamiętać?	156
Zestaw do kompresji Squeeze	156
Jak wygląda ten program?	159
W jaki sposób działa ten program?	161
Czym wyróżnia się ten program?	166
Na co należy zwrócić uwagę podczas używania tego programu?	166
Co powinienś zapamiętać?	166
Episode Pro	167
Jak wygląda ten program?	168
W jaki sposób działa ten program?	169
Czym wyróżnia się ten program?	173
Na co należy zwrócić uwagę podczas używania tego programu?	174
Co powinienś zapamiętać?	174
Rozwiązania wyższej i niższej klasy	174
Rozwiązania tańsze	175
Uwagi na temat rozwiązań dla rozległych środowisk	177
Zakończenie	178

Wywiad z kompresjonistą: Nico Puertollano 179

Rozdział 6. Kompresowanie dla DVD 183

Co to jest DVD?	184
Tworzenie płyt DVD o standardowej jakości (SD)	185
Media standardowej jakości.	185
Kodeki dla DVD o standardowej jakości.	188
A cóż to takiego VOB?	189
Co to jest dysk Blu-ray?	190
Tworzenie dysków Blu-ray	193
Przegląd technologii Blu-ray	193
Nagrywanie dysków Blu-ray	196
Sposoby tworzenia DVD	197
Sprawa ustalania optymalnej prędkości transmisji danych	200
Metody tworzenia płyt DVD o standardowej rozdzielczości	202
Kompresja i markery w programie DVD Studio Pro	202
Generowanie pliku zgodnego z DVD w aplikacji Compressor	204
Tworzenie płyty SD DVD w programie DVD Studio Pro	211
Metody tworzenia płyt Blu-ray	217
Praca z programem Premiere Pro	218
Zarządzanie procesem przekazywania (handoff)	218
Generowanie pliku w formacie MPEG-2 zgodnego ze standardem Blu-ray w aplikacji Adobe Media Encoder	219
Generowanie pliku w formacie H.264 zgodnego ze standardem Blu-ray w aplikacji Adobe Media Encoder	222
Tworzenie płyty DVD w standardzie Blu-ray w programie Encore	225
Zakończenie	229

Wywiad z kompresjonistą: Ben Waggoner 231

Rozdział 7. Kompresowanie dla Internetu 237

Wczesne lata internetowego wideo	238
Nowoczesne wideo w Internecie	240
Wzrost zainteresowania krótkotrwałymi treściami wideo	240
Powszechność łączy szerokopasmowych	241
Lepsze narzędzia	241
Widz i twórca to te same osoby	243
Każdy może być gwiazdą	243
Dziennikarstwo obywatelskie	244
Tradycyjne media przystosowują się	244
Nowoczesne narzędzia i formaty sieciowe	245

Metody kompresowania wideo dla Internetu	247
Kodowanie dla YouTube przy użyciu aplikacji Episode Pro	247
Klip strumieniowy w formacie Windows Media stworzony w aplikacji Compressor	255
Plik wideo w formacie Flash Video do pobierania progresywnego	264
Zakończenie	266

Wywiad z kompresjonistą: Jim Rohner 267

Rozdział 8. Kompresowanie dla urządzeń przenośnych 271

W jaki sposób nieduże wideo stało się wielkim	272
Transmisje strumieniowe na żywo do urządzeń przenośnych	273
Pobieranie i (być może) synchronizowanie	274
Urządzenia przenośne i ich popularne formaty	278
Telefony komórkowe	278
Telefony inteligentne i komputery Pocket PC	281
iPhone firmy Apple	282
Odtwarzacze mediów	284
Sposoby kompresji plików wideo dla urządzeń przenośnych	290
Pliki 3GP dla telefonów, które nie są inteligentne	290
Dodawanie znaczników rozdziałów do podkastu wideo	301
Zakończenie	308

Wywiad z kompresjonistką: Rynne Hodson 309

Rozdział 9. Kompresowanie dla urządzeń set-top box..... 313

Dużo więcej niż zwykły kabel	314
Telewizja cyfrowa	315
IPTV	316
Urządzenia STB oparte na protokole IP	318
Apple TV	318
Xbox 360	321
Vudu	323
Sposoby kompresji wideo dla urządzeń set-top box	324
Ripowanie płyt DVD przy użyciu programu HandBrake	325
Plik wideo o standardowej rozdzielczości dla urządzenia Xbox	333
Zakończenie	336

Wywiad z kompresjonistą: Andy Beach 337

Skorowidz 341



Zrozumieć technologię wideo i audio

Kompresja wideo jest jednym z kluczowych elementów świata produkcji. Jest ona codziennie cichym wybawcą (i mordercą) wielu projektów. Jest jednak procesem niezauważalnym, gdy wykonuje się ją w sposób prawidłowy. Kiedy ostatnio podczas oglądania programu telewizyjnego zachwycałeś się wspaniałą jakością kompresji? I przeciwnie, zła kompresja jest od razu widoczna i może zepsuć prawie każde nagranie wideo.

Podstawowe zadanie kompresji wideo jest całkiem proste: dostosować dane wideo do żądanego sposobu dostarczania — telewizji, DVD, Internetu, Twojego iPod'a lub telefonu komórkowego. Trudnością staje się takie dopasowanie działania w zakresie wymaganych specyfikacji technicznych i ograniczeń wymuszanych przez medium transportowe, aby dostarczyć doznania audiowizualne, które będą satysfakcjonować użytkownika końcowego. Na przykład, jeśli udostępniasz zawartość wideo w Internecie, będziesz musiał uwzględnić problemy związane z rozmiarem danych. Mógłbyś mieć najwspanialszy film wszech czasów, lecz jeśli byłby on zbyt duży i nikt nie mógłby go pobrać, kto by go właściwie oglądał? Podobnie jeśli Twoje dane wideo będą nadawane w telewizji, musisz mieć gwarancję, że wszystkie pola i klatki przetworzonego i skompresowanego programu będą wciąż nienaruszone po długotrwałym procesie tworzenia.

Ale teraz wybiegam chyba za bardzo do przodu. Sprawy przedstawiają się w ten sposób: istnieje pewna podstawowa wiedza, którą trzeba znać, zanim zagłębimy się w złożoności związane z kompresją wideo. Jeśli już od dłuższego czasu zajmujesz się techniką wideo, wszystkie informacje są Ci na pewno dobrze znane.

W rzeczywistości nawet nie myślisz o nich w sposób świadomy podczas codziennej pracy. Jeśli jednak jesteś dopiero początkujący w dziedzinie wideo, zapoznanie się z podstawami jej działania pozwoli Ci o wiele szybciej rozwiązywać kolejne problemy związane z kompresją.

Dziękujemy Ci, Philo T. Farnsworthie

Pierwszy pokaz techniki, dzięki której powstała nowoczesna telewizja i wideo, odbył się 27 września 1927 roku w San Francisco dzięki staraniom Philo T. Farnswortha. W wieku lat 14 wymyślił on samą ideę w Rigby, Idaho.

Istniało oczywiście wiele innych osób i korporacji, które zajmowały się tworzeniem i usprawnianiem elektronicznej telewizji Farnswortha, lecz podstawowy pomysł wideo powstał w umyśle tego trochę zapomnianego młodego wynalazcy. Idee Farnswortha są wciąż podstawą technologii wideo w obecnych czasach mimo znacznego rozwoju techniki, który miał miejsce w ciągu ostatnich 80 lat.

Garść informacji z życia Philo

- ▶ Pierwszym wysłanym obrazem był znak dolara.
- ▶ Transmisja składała się z 60 poziomych linii.
- ▶ Farnsworth opracował lampę analizującą, będącą podstawą wszystkich telewizorów kineskopowych.
- ▶ Uzyskał on pierwszy patent dla swojej lampy analizującej, natomiast późniejszy patent utracił na rzecz RCA.
- ▶ Wynalazł ponad 165 różnych urządzeń, między innymi urządzenie służące do zamiany obrazu na sygnał elektryczny, wzmacniacz, promień katodowy, lampy próżniowe, skanery elektryczne, powielacze elektronów oraz materiały fotoelektryczne.

Elementy wideo

W przeciwieństwie do filmu, którego działanie polega na wyświetlaniu obrazów, **wideo** jest sygnałem elektronicznym. Chociaż termin *wideo* był pierwotnie używany w celu oznaczenia dowolnego sygnału nadawanego (lub *teletransmitowanego*) do odbiornika telewizyjnego, to jednak z upływem czasu został tak przededefiniowany, aby opisywać obrazy wyświetlane w sposób elektroniczny (takie, które są używane w reklamowych tablicach wideo, telefonach komórkowych, bankomatach itd.).

Wideo stało się wszechobecną częścią naszego życia, szczególnie od momentu, gdy pojawiły się komputery (i ich monitory). Używamy techniki wideo podczas współpracy z innymi ludźmi, dla rozrywki, w celach komunikacyjnych, ale również w takich przypadkach jak zwykle pobranie gotówki z banku. Technologia zmieniła tak wiele, że obecnie w wieczornych wiadomościach często oglądamy fragmenty filmów wykonanych za pomocą telefonów komórkowych. Zmieniły się zarówno sposoby wykorzystania wideo, jak również wspierająca je technologia.

Klatki i pola

Gdy zbiór sekwencyjnych obrazów zostaje szybko wyświetlony, pojawia się interesujące zjawisko. Zamiast odbioru kolejnych nieruchomych obrazów ludzie widzą płynnie zachodzącą animację. Zjawisko to znane jest jako **bezwładność wzroku** i jest podstawą działania filmu i wideo. Liczba obrazów wyświetlanych w ciągu sekundy zwana jest **szybkością klatek** (sekundy są najczęściej stosowaną miarą szybkości klatek, lecz nie jedyną). Już od szybkości równej około 8 *klatek na sekundę* (w skrócie *fps*) obserwator zaczyna widzieć płynny ruch; jednakże zauważa on także wyraźne migotanie i nieregularne odtwarzanie. Aby ustrzec się przed takim migotaniem między kolejnymi klatkami, potrzeba generalnie szybkości większej niż 16 klatek na sekundę (choć jest to opinia subiektywna, a wielu ludzi uważa, że w celu całkowitego pozbycia się migotania wymagane jest zapewnienie prędkości 24 – 30 fps). Im szybsze ma być wrażenie ruchu, tym więcej klatek potrzebujesz, aby uzyskać odpowiednią jego płynność. Nowoczesny film ma szybkość klatek równą 24 fps, natomiast w przypadku telewizji parametr ten wynosi 25 fps w Polsce i innych krajach używających standardu PAL (ang. *Phase Alternating Line*). Tę samą prędkość wykorzystuje inny standard SECAM (franc. *Sequentiel Couleur Avec Memoire*), używany przeważnie w krajach frankońskich i Rosji. Wreszcie standard NTSC (ang. *National Television Standards Committee*) stosuje szybkość klatek równą około 30 fps (dokładnie 29,97 fps). Jest on używany między innymi w Stanach Zjednoczonych i Japonii.

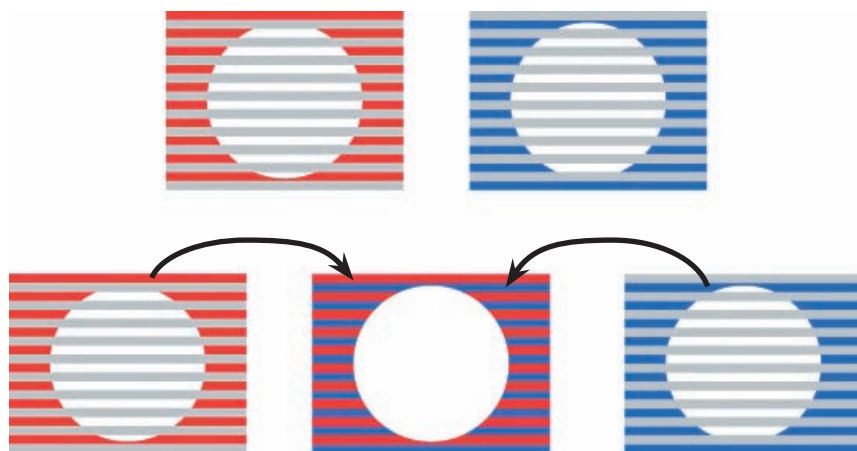
Klatka może zostać zaprezentowana widzowi w dwojaki sposób: przy użyciu **skanowania progresywnego** (ang. *progressive scanning*) lub **skanowania z przeplotem** (ang. *interlaced scanning*). W dzisiejszych czasach jesteś na pewno tego bardziej świadomy niż kiedykolwiek wcześniej, gdyż dążymy do telewizji wysokiej rozdzielczości (HDTV). W przypadku parametrów specyfikacji technicznej telewizji HDTV, takich jak 1080i czy 720p, litery *i* oraz *p* oznaczają odpowiednio skanowanie z przeplotem oraz skanowanie progresywne (wartości 1080 oraz 720 określają liczbę linii w pionie; więcej na ten temat w dalszej części książki).

Skanowanie z przeplotem zostało opracowane we wczesnych latach trzydziestych jako sposób na poprawę wyświetlania obrazów w monitorach wyposażonych w lampę kineskopową (CRT — ang. *Cathode Ray Tube*). W takie lampy wyposażone były wszystkie telewizory aż do momentu, gdy pojawiły się ekrany plazmowe

i LCD. W lampie kineskopowej porusza się wiązka elektronów, uderzając w powierzchnię ekranu pokrytego od wewnątrz luminoforem będącym związkiem chemicznym emitującym światło. Dawniej te związki chemiczne miały bardzo krótki czas poświaty (obecne kineskopowe monitory komputerowe mają dłuższy czas poświaty). Gdy wiązka elektronów zbliżała się do dolnych linii ekranu, luminofor w górnej części był już ciemny. Aby rozwiązać ten problem, dawni inżynierowie telewizji zaprojektowali system przeplotu służący do skanowania wiązką elektronów. Za pomocą systemu przeplotu wiązka skanuje najpierw wyłącznie nieparzyste linie ekranu, posuwając się od góry w dół ekranu, a następnie powraca do punktu wyjściowego i rozpoczyna skanowanie linii parzystych. Te dwie rozłączne grupy linii (pokazane na rysunku 1.1) znane są w przypadku sygnału telewizyjnego jako *pole górne* (lub *nieparzyste*) i *pole dolne* (lub *parzyste*). Tak więc system telewizyjny, działający z szybkością 30 fps, czyni to w rzeczywistości z prędkością 60 fps, wyświetlając dwa obrazy z przeplotem w ramach jednej klatki.

Rysunek 1.1.

System wideo z przeplotem skanuje najpierw nieparzyste, a następnie parzyste linie, łącząc ostatecznie pola, aby uzyskać kompletną klatkę. Używa on obu pól z przeplotem, aby wygenerować pełny obraz



Korzyści ze stosowania przeplotu

Wszystkie systemy wideo zawierają w sobie rozwiązania kompromisowe, takie jak zależność między jakością obrazu a rozmiarem pliku lub liczbą bitów przeznaczonych na dane wideo w porównaniu liczbą bitów wykorzystywanych w ścieżce audio. Jednym z najważniejszych czynników, które należy uwzględnić, jest **szerokość pasma** (ang. *bandwidth*), określana w megahercach w przypadku wideo analogowego lub za pomocą **szybkości transmisji** (ang. *bit rate*) dla wideo cyfrowego. Im większa jest szerokość pasma, tym bardziej złożona i kosztowna musi być cała infrastruktura, poczynając od kamery, której używa się do nagrywania wideo, a kończąc na monitorach, na których wyświetlone zostaną obrazy po ich przesłaniu. Dla określonej liczby linii oraz wartości odświeżania zastosowanie systemu wideo z przeplotem redukuje szerokość pasma sygnałowego o współczynnik równy 2.

Uwaga: Opisywanie sygnału telewizyjnego za pomocą szerokości pasma może wydawać się dziwne, lecz jest to dokładnie to, czego potrzebujemy. W przypadku standardowej telewizji szerokość pasma mierzona jest w hercach, natomiast dla plików wideo, umieszczanych w Internecie, szerokość pasma podawana jest w bitach.

Większa wartość szerokości pasma może umożliwić dostarczenie sygnału wideo z przeplotem o dwukrotnie wyższej wartości odświeżania w porównaniu z systemem o skanowaniu progresywnym. Powoduje to redukcję migotania na monitorach kineskopowych. Większa szybkość transmisji poprawia odtwarzanie ruchu, gdyż pozycja ruchomego obiektu jest wizualizowana i uaktualniana dużo częściej.

Problemy podczas stosowania przeplotu

Istnieją oczywiście wady technologii wideo z przeplotem. Zaprojektowana została ona w taki sposób, aby obrazy były nagrywane, przesyłane, zapisywane i wyświetlane w tym samym formacie z przeplotem. Ponieważ każda klatka wideo z przeplotem składa się z dwóch pól, które są nagrywane w różnych momentach czasowych, obrazy zapisane w taki sposób będą zawierać *artefakty* (sztuczne wady) ruchu, gdy oba pola zostaną połączone i wyświetlone w tym samym czasie.

Ogólnie rzecz biorąc, format z przeplotem zastępowany jest stopniowo przez wideo progresywne (w przypadku którego każdy przechwycony obraz zapisywany jest w postaci pełnej klatki zamiast dwóch pól). Nawet kamery wideo mogące nagrywać jedynie obrazy z przeplotem zaczynają być wyposażane w pseudo-progresywne tryby, które pozwalają na tworzenie złożonych klatek o czasie odświeżania równym połowie standardowej wartości. Wszystkie nowoczesne komputery używają skanowania progresywnego w celu wyświetlania obrazu, a nowsze rodzaje ekranów wideo, takie jak LCD i plazma, są specjalnie zaprojektowane jako monitory ze skanowaniem progresywnym, wyświetlające wszystkie poziome linie obrazu w każdej klatce wideo. Gdyby te monitory wyświetlały wideo z przeplotem, wynikiem byłby obraz, który miałby ograniczoną rozdzielczość w poziomie oraz artefakty ruchu jak te zaprezentowane na rysunku 1.2. Te artefakty mogą być również widoczne, gdy wideo z przeplotem jest wyświetlane z prędkością niższą niż podczas nagrywania. Przykładem tego jest odtwarzanie wideo w zwolnionym tempie.

Ponieważ nowoczesne komputerowe ekrany wideo są systemami ze skanowaniem progresywnym, wideo z przeplotem będzie generowało widoczne artefakty podczas wyświetlania na nich. Obecnie większość edycji wideo przeprowadzana jest w komputerze, dlatego też ta rozbieżność między komputerowymi systemami



Rysunek 1.2. Po lewej stronie obraz z przeplotem (dwa pola); po prawej stronie ta sama klatka po usunięciu przeplotu

służącymi do wyświetlania wideo a formatami sygnałów telewizyjnych oznacza, że edytowana zawartość wideo nie może być właściwie prezentowana, dopóki nie użyje się dodatkowego urządzenia, takiego jak monitor działający w standardzie PAL. Na szczęście z tego powodu większość systemów oraz profesjonalne oprogramowanie służące do edycji wideo wspierają możliwość podglądu obrazu na oddzielnych monitorach.

Aby zminimalizować powstawanie artefaktów generowanych przez wideo z przeplotem na monitorze ze skanowaniem progresywnym, większość rozwiązań ma możliwość stosowania filtrów usuwających przeplot. **Usuwanie przeplotu** (ang. *deinterlacing*) jest procesem polegającym na przekształcaniu pól z przeplotem na sekwencję klatek bez przeplotu, tak jak przedstawiono na rysunku 1.2. Nie jest to proces idealny i generalnie kończy się utworzeniem obrazów o mniejszej rozdzielczości, szczególnie w miejscach, gdzie pojawiają się obiekty w ruchu. Procedura ta wymaga standardowo użycia tzw. **tymczasowej interpolacji** (ang. *temporal interpolation*). Nazwa ta jest ozdobnym określeniem próby odgadnięcia, w którym miejscu wideo pojawi się ruch obiektów, a następnie wykonania takich operacji łączenia i korekcji tego ruchu, aby usunąć zakłócenia występujące na przetwarzanych klatkach. Systemy usuwające przeplot są wbudowane w monitory telewizyjne działające na zasadzie skanowania z progresją. Pozwala to na osiągnięcie najwyższej jakości wyświetlanych obrazów w przypadku sygnałów wideo z przeplotem.

Przeplot wprowadza również możliwość wystąpienia problemu zwanego **migotaniem międzyliniowym** (ang. *interline twitter*) lub **morą** (ang. *moiring*). Ten efekt interferencyjny pojawia się jedynie w określonych okolicznościach, gdy obraz zawiera pewne pionowe elementy mające rozmiary zbliżone do poziomej rozdzielczości danego formatu wideo. Na przykład, osoba występująca w telewizji, ubrana w koszulę mającą delikatny wzór złożony z ciemnych i jasnych linii, może zostać przedstawiona na ekranie w taki sposób, jakby pasy na jej koszuli „migotały”. Profesjonaliści telewizyjni wiedzą, że nie powinno nosić się ubrań o takich wzorach, które mogłyby powodować problemy z interferencją. W kamerach wideo wysokiej klasy można użyć specjalnych filtrów, które zastosowane dla rozdzielczości pionowej sygnału uchronią obraz przed pojawieniem się mory.

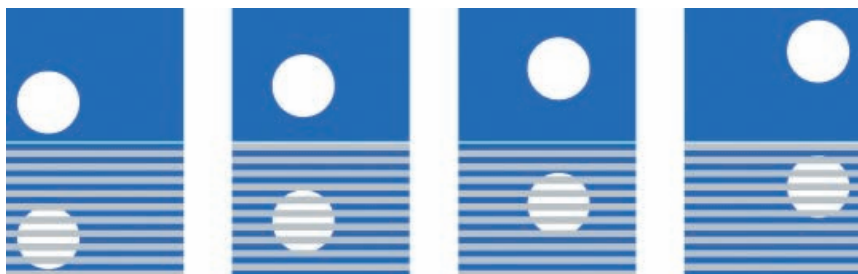
Wreszcie wideo dostarczane przez Internet, zarówno poprzez przesyłanie strumieniowe, jak i zwykłe pobieranie plików (więcej na temat metod transportu w rozdziale 2.), jest skompresowane i udostępnione w postaci progresywnej. Istnieją różne opinie na temat tego, czy koniecznie powinno się od razu nagrywać w formacie progresywnym, czy też wystarczy zastosować usuwanie przeplotu, aby stworzyć efektywne wideo dla Internetu. Ważne jest natomiast jedno: wideo musi mieć już format progresywny przed dostarczeniem go przez Internet i urzędu przeniósł.

Wideo ze skanowaniem progresywnym

W przypadku skanowania progresywnego cała klatka wideo zostaje przechwycona w jednym kroku, zamiast tworzyć dwa osobne pola z przeplotem. Skanowanie progresywne ma wiele zalet w porównaniu ze skanowaniem z przeplotem, wśród których wyliczyć można brak migotania międzyliniowego oraz innych problemów. Aby jednakże dostarczyć ten sam poziom płynności ruchu, wymagana jest dwukrotnie większa szerokość pasma niż w przypadku transmisji sygnału z przeplotem (choć podczas odtwarzania obraz w formacie progresywnym będzie miał także dwukrotnie większą rozdzielczość niż obraz z przeplotem). Rysunek 1.3 przedstawia różnicę między klatkami z przeplotem a klatkami progresywnymi.

Rysunek 1.3.

Zwróć uwagę, o ile płynniejszy jest ruch na obrazie z progresją (na górze rysunku) w porównaniu z obrazem z przeplotem (na dole rysunku)



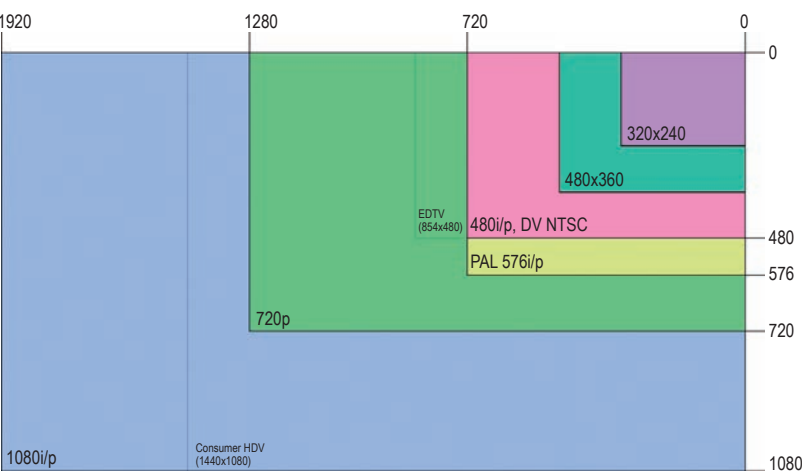
Rysunek 1.3 przedstawia poruszającą się piłkę. Obraz z przeplotem (na dole rysunku) musi wyświetlić pola, które są trochę niesynchronizowane ze sobą (ponieważ piłka ciągle się porusza). Może powodować to pewne zniekształcenia lub pogorszenie jakości obrazu. I przeciwnie — obraz progresywny (na górze rysunku) wyświetla za każdym razem kompletną klatkę, dlatego też jego jakość jest lepsza, a ruch bardziej płynny, mimo że w celu transmisji konieczne jest użycie pasma o większej szerokości.

Z wielu powodów wideo progresywne zastąpiło wideo z przeplotem w różnych obszarach świata multimedialnych. Jednym z tych powodów jest to, iż nowoczesne technologie telewizyjne są w zasadzie progresywne. Wielu producentów zaprzestało produkcji tradycyjnych telewizorów kineskopowych, a rynek zdobywany jest przez nowsze technologie wyświetlania obrazu, takie jak Digital Light Projection (DLP), telewizory plazmowe i LCD.

Technologie transmisji danych również ulegają zmianie. Gdy telewizja używała jedynie sygnałów analogowych, wideo z przeplotem pozwalało na uzyskanie wąskopasmowej transmisji umożliwiającej dostarczanie obrazów o odpowiedniej jakości i płynności ruchu. Lecz obecnie analogowy system nadawania zastępowany jest przez transmisję cyfrową oraz cyfrowy sposób dostarczania danych (przez Internet, satelity, sieci kablowe oraz dyski optyczne), dla których wideo progresywne jest bardziej efektywne. Dodatkowo ciąg czynności technologicznych, uwzględniający standard skanowania progresywnego (który pojawił się najpierw w monitorach komputerowych i wysokiej klasy ekranach ściennych), ma sens, odkąd zaczęto stosować go w odbiornikach telewizyjnych ogólnego użytku przeznaczonych dla szerokiego kręgu konsumentów.

Rozdzielczości

Jakość obrazów, które widzisz na filmie lub wideo, nie jest jedynie funkcją liczby klatek wyświetlanych w ciągu sekundy lub metody ich przedstawiania (pełne klatki z progresją lub pola z przeplotem). Ważnym czynnikiem jest również ilość informacji w każdej klatce, zwana **rozdzielczością obrazu** (ang. *image resolution*). Na rysunku 1.4 możesz zauważyć, że rozdzielczość obrazu zmienia się znacznie w zależności od rodzaju ekranu. Standardowa telewizja PAL zajmuje obszar różowo-żółty (rozdzielczość 720 na 576), podczas gdy dla nowoczesnej telewizji wysokiej rozdzielczości przeznaczono dwa większe pola: 1080p (o rozdzielczości 1920 na 1080) oraz 720p (o rozdzielczości 1280 na 720).



Rysunek 1.4. Oto niektóre z najbardziej popularnych rozdzielczości obrazów wideo używanych w telewizji, odtwarzaczach DVD, komputerach, Internecie oraz urządzeniach przenośnych. Liczby na górze rysunku oznaczają szerokość obrazu wyrażoną w pikselach. Liczby po prawej stronie rysunku określają wysokość obrazu, również w pikselach

Rozdzielczość wideo analogowego reprezentowana jest przez liczbę skanowanych linii przypadających na dany obraz, która w rzeczywistości oznacza liczbę linii rysowanych przez wiązkę elektronów w czasie, gdy przesuwa się ona z góry na dół ekranu.

Rozdzielczość obrazów cyfrowych, wyświetlanych przykładowo na monitorach komputerowych oraz cyfrowych telewizorach, jest określana przez stałą liczbę podstawowych elementów ekranu (pikseli) i jest często wyrażana jako rozmiar: liczba pikseli w poziomie oraz liczba pikseli w pionie. Na przykład, 640 na 480 oraz 720 na 480 oznaczają rozdzielczości pełnej klatki dla standardowej jakości obrazu, natomiast 1920 na 1080 określa rozdzielczość pełnej klatki dla wysokiej jakości obrazu.

Uwaga: Dlaczego istnieją dwa rodzaje rozdzielczości dla standardowej jakości obrazu? Zależy to od współczynnika kształtu piksela, o którym opowiemy w rozdziale 2. (poszukaj podrozdziału zatytułowanego „Piksele o różnych kształtach”).

Rozdzielczość pionowa dla standardów NTSC i PAL

Format NTSC oparty jest na rozdzielczości wynoszącej 525 poziomych linii wyświetlanych jako dwa pola z przeplotem. Jednakże niektóre z tych linii są używane do celów synchronizacji i wygaszania, dlatego też w rzeczywistości na *aktywnym* obszarze ekranu widocznych jest jedynie 486 linii. (Wszystkie standardy wideo — nie tylko NTSC — zawierają więcej informacji niż tylko wyświetlana zawartość klatki. Przed i po obrazie znajdują się linie i piksele, które zawierają informację o synchronizacji oraz opóźnienie czasowe. Ten otaczający margines zwany jest **obszarem wygaszania** [ang. *blanking interval*].)

Format PAL jest oparty na rozdzielczości wynoszącej 625 poziomych linii wyświetlanych jako dwa pola z przeplotem. Podobnie jak w przypadku NTSC niektóre z tych linii są używane do celów synchronizacji i wygaszania, dlatego też na aktywnym obszarze ekranu widocznych jest faktycznie 576 linii.

Typowe rozdzielczości dla telewizji w standardzie PAL

576i – 576 linii w rozdzielczości pionowej, skanowanych z przeplotem.

576p – 576 linii w rozdzielczości pionowej, skanowanych progresywnie.

1080i – 1080 linii w rozdzielczości pionowej, skanowanych z przeplotem.

Współczynnik kształtu

Stosunek szerokości obrazu do jego wysokości zwany jest **współczynnikiem kształtu** (ang. *aspect ratio*). Zachowanie prawidłowego współczynnika kształtu obrazu jest jednym z istotniejszych elementów kompresji wideo. W trakcie skalowania wideo do odpowiedniego rozmiaru, aby dopasować je do różnych ekranów i rozdzielczości, można łatwo stracić zależność między oryginalną wysokością i szerokością obrazu. Gdy to się przydarzy, można uzyskać zniekształcony obraz, czasem nawet niemożliwy do oglądania.

Klatki błony fotograficznej 35 mm, na której budowie opierała się pierwotnie taśma filmowa, miały współczynnik kształtu 4:3 (stosunek szerokości do wysokości). Jest on również oznaczany jako współczynnik 1,33:1 lub 1,33 (mnożąc wysokość przez 1,33, uzyskuje się szerokość).

Od samego początku powstania przemysłu filmowego aż do wczesnych lat 50. współczynnik kształtu 4:3 był używany prawie wyłącznie podczas tworzenia filmów i określania wyglądu ekranów kin. Po uruchomieniu telewizji istniejące soczewki kamer były przystosowane wyłącznie do formatu 4:3, dlatego też ten sam współczynnik kształtu został wybrany jako standard dla nowego medium przesyłania informacji. Obecnie format 4:3 oznacza **telewizję pełnoekranową** (ang. *full-screen TV*).

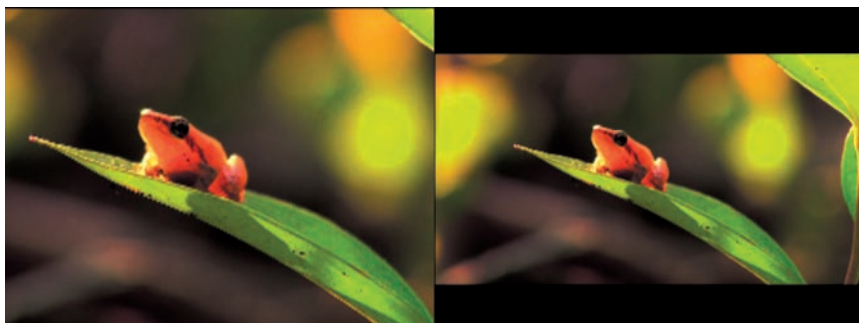
W latach 50. przemysł filmowy zauważył zależność między rosnącymi wpływami telewizji a malejącą liczbą widzów w kinach. Dlatego też studia filmowe wprowadziły szereg ulepszeń, umożliwiając widzom przeżycie bardziej ekscytujących doznań, niż byłoby to możliwe w przypadku oglądania telewizji w mieszkaniach. Najbardziej znaczącym ulepszeniem było upowszechnienie szerokiego ekranu. Studia produkowały filmy panoramiczne w różnych formatach „scope”, na przykład Cinemascope (format pierwotny), Warnerscope, Techniscope i Panascope.

Podstawowym problemem takich formatów panoramicznych jest to, że nie mogą zostać poprawnie skonwertowane na obszar ekranu telewizyjnego. Gdy na zwykłym telewizorze wyświetlane są filmy panoramiczne, boczne fragmenty obrazu są najczęściej obcięte, aby dopasować się do współczynnika kształtu 4:3, jak przedstawiono po lewej stronie rysunku 1.5. Proces ten nazywany jest **pan & scan**, ponieważ wyboru interesującego fragmentu obrazu dokonuje się dynamicznie, biorąc pod uwagę akcję rozgrywającą się aktualnie na ekranie. Aby tego uniknąć, studia często używają techniki o nazwie **letterbox** — czarnych pasów umieszczonych ponad i poniżej obrazu panoramicznego, aby go wiernie odwzorować, jak przedstawiono po prawej stronie rysunku 1.5.

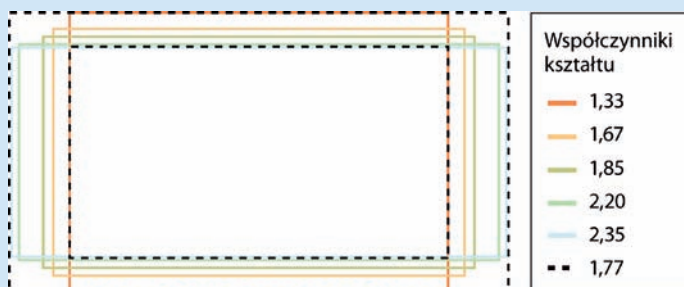
Wprowadzanie telewizji o wysokiej rozdzielczości wymusza przejście ze standardowego formatu telewizyjnego 4:3 na nowszy format panoramiczny. Współczynnik kształtu telewizji panoramicznej wynosi 16:9 (1,78), co całkiem dobrze odpowiada najbardziej popularnemu współczynnikowi dla filmów kinowych równemu 1,85. W tabeli 1.1. przedstawiono aktualne i pojawiające się formaty wideo, jak również ich współczynniki kształtu, rozdzielczości, szybkości klatek oraz szybkości transmisji (tam, gdzie to było możliwe).

Rysunek 1.5.

Po lewej stronie obraz panoramiczny, skadrowany do formatu 4:3. Po prawej stronie ten sam obraz wyświetlony w całości dzięki użyciu techniki letterbox

**Dlaczego 16:9?**

Dr Kerns Powers z Centrum Badawczego Davida Sanroffa w Princeton, New Jersey, będącego czołowym laboratorium rozwojowym technologii telewizyjnych, przeanalizował wszystkie najczęściej stosowane współczynniki kształtu, a następnie zestawiał je razem. Wówczas odkrył coś interesującego. Gdy wziął pod uwagę prostokąt o określonych proporcjach i przeskalował go za pomocą dwóch różnych metod, mógł zawrzeć w nim inne prostokąty utworzone przy użyciu wszystkich współczynników kształtu branych pod uwagę. Ten magiczny prostokąt miał proporcje równe szesnastu jednostkom długości na dziewięć jednostek szerokości, inaczej mówiąc 16:9 (rysunek 1.6). Dzięki temu odkryciu format 16:9 stał się nowym standardem współczynnika kształtu dla telewizji HDTV, a większość telewizorów wysokiej rozdzielczości również ma ekrany o proporcjach 16:9.



Rysunek 1.6. Sposób rozwiązania problemu współczynników kształtu zaproponowany przez Kernsa polegał na zebraniu wszystkich często używanych rozdzielczości i nałożeniu ich na siebie. Po wykonaniu tej czynności odkrył on, że wszystkie te elementy można by zawrzeć w prostokącie o proporcjach 16:9

Większość sprzętu produkcyjnego lub postprodukcyjnego, który jest obecnie używany, pozwala na nagrywanie i odtwarzanie wideo w tych formatach. Nowe medialne urządzenia wideo, takie jak telefony komórkowe i strony internetowe, mają zupełnie inne wymagania dotyczące formatów, które nie mogą zostać ujednolicone za pomocą standardowych specyfikacji.

30 Kompresja dźwięku i obrazu wideo

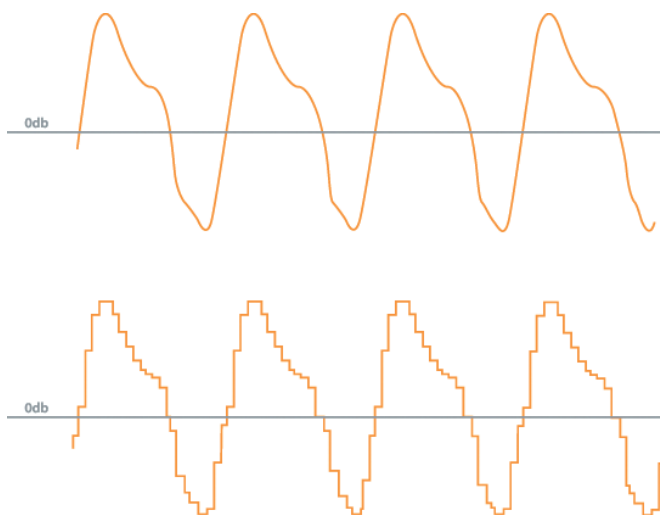
Tabela 1.1. Nowoczesne formaty wideo i ich współczynniki kształtu

Format	Współczynnik kształtu	Rozdzielczość pozioma (piksele/linie)	Rozdzielczość pionowa (linie skanowania)	Szybkość klatek	Szybkość transmisji (megabity/sekundę)
NTSC (Stany Zjednoczone, Kanada, Japonia, Korea, Meksyk)	4:3	330	525 (480 linii widocznych)	30i	—
PAL (Australia, Chiny, większość krajów Europy, Ameryka Południowa)	4:3	330	625 (576 linii widocznych)	25i	—
SECAM (Francja, Bliski Wschód, większość krajów Afryki)	4:3	330	625 (576 linii widocznych)	25i	—
				24p	18 Mb/s
HDTV	16:9	1920	1080	30p	18 Mb/s
				30i	18 Mb/s
				24p	8 Mb/s
HDTV	16:9	1280	720	30p	10 Mb/s
				60p	18 Mb/s
				24p	3 Mb/s
SDTV	16:9	720	483	30p, 30i	4 Mb/s, 4 Mb/s
				60p	8 Mb/s
				24p	3 Mb/s
SDTV	4:3	720	486	30p, 30i	4 Mb/s, 4 Mb/s
				60p	7 Mb/s
				24p	3 Mb/s
SDTV	4:3	640	480	30p, 30i	3 Mb/s, 3 Mb/s
				60p	7 Mb/s

Sygnały analogowe a cyfrowe

Jak wcześniej wspomniano, sygnały wideo mogą być analogowe lub cyfrowe. Sygnał analogowy (rysunek 1.7) jest wciąż zmieniającym się napięciem elektrycznym, który dla danego przedziału czasu może zostać przedstawiony w postaci fali. Poszczególne linie obrazu wideo są zapisywane razem z informacją o czasie, aby umożliwić odbiornikowi (czyli Twojemu telewizorowi) poprawne odtworzenie informacji.

Rysunek 1.7.
Oto najprostsze
porównanie
sygnału analogowe-
go (góra rysunku)
z cyfrowym
(dół rysunku)



Z drugiej strony, sygnał cyfrowy jest liczbową reprezentacją sygnału analogowego. Oznacza to, że sygnał cyfrowy jest w rzeczywistości strumieniem **bitów** (danymi przechowywanymi w postaci długiego ciągu liczb binarnych). Każda liczba znajdująca się na liście jest zdjęciem (inaczej zwanym **próbką** — ang. *sample*) sygnału analogowego w dowolnym momencie czasu. Prędkość próbkowania strumienia cyfrowego jest liczbą zdjęć wykonanych w ciągu sekundy.

Sygnały cyfrowe mają wiele zalet w porównaniu z analogowymi. Jedną z najważniejszych jest jakość odtwarzania nagrania, zwana **wiernością** (ang. *fidelity*). Urządzenie analogowe, takie jak magnetowid, przekształca w prosty sposób zmiany napięcia na dźwięk lub obraz, lecz nie może odróżnić oryginalnego sygnału od napięcia pochodzącego z zakłócenia elektrycznego (takiego jak sieć elektryczna). Zakłócenie elektryczne może pochodzić z zewnątrz, z taśmy lub z komponentów samego magnetowidu lub telewizora. Gdy wykonujesz operację kopiowania, zakłócenia nagrane na nośniku źródłowym przenoszą się na nową taśmę. Jeśli musiałbyś następnie skopiować tę nową taśmę, zakłócenia z dwóch poprzednich taśm pojawiłyby się na trzeciej taśmie i tak dalej; każda wykonana kopia pogarsza wierność odtwarzania oryginału. Efekt ten zwany jest **pokoleniową degeneracją** (ang. *generation loss*).

W przypadku techniki cyfrowej sygnał nagrywany na taśmę składa się wyłącznie z łańcuchów zer i jedynek, które zostają następnie przekształcone przez odtwarzacz cyfrowy na liczby, a w końcu na obrazy lub dźwięki. Ponieważ odtwarzacz cyfrowy potrafi czytać tylko zera i jedynki, może on dużo łatwiej rozpoznać oryginalny sygnał i zakłócenia. Dlatego też możesz przysyłać i kopiować sygnały cyfrowe tak często, jak tylko chcesz — bez pogorszenia lub w najgorszym razie z minimalnym pogorszeniem wierności.

W jaki sposób działa kompresja

Działanie kompresji wideo można najprościej zdefiniować jako analizowanie wartości każdej klatki i ustalanie, w jaki sposób można ją odtworzyć przy użyciu mniejszej ilości informacji (technologicznym równoważniku parafrazowania). Jest to możliwe dzięki stosowaniu **kodeków** (ang. *codecs*). Nazwa ta jest skrótem od słów „algorytm *kompresji/dekompresji*”. Kodeki wykonują swoje zadanie w różny sposób. Powiedzmy, że mamy całkowicie czarną klatkę o jakości cyfrowej (przed pojawieniem się Twoich danych); kodek musi pamiętać tylko o jednym: „każdy piksel tej klatki ma taki sam odcień koloru czarnego”. Jest to dużo mniej informacji niż podawanie ciągu liczb „0, 0, 0” 345 600 razy (co dla tych Czytelników, którzy za mną nadążają, jest równe liczbie 720 pomnożonej przez 480).

Ale przecież większość obrazów wideo nie składa się tylko z jednego koloru! Dlatego też kodek musi odnaleźć miejsca, w których klatki wartości danych różnią się między sobą — na przykład wykryć granice między jasnymi i ciemnymi elementami — a następnie zapisać te informacje w sposób bardziej efektywny. Wykonywane jest to poprzez podział sceny na grupy pikseli zwane **makroblokami** (ang. *macroblocks*) oraz reprezentowanie ich za pomocą liczb, dzięki którym można następnie odtworzyć elementy obrazu (na takiej zasadzie działają kodeki wykorzystujące dyskretną transformatę kosinusową, takie jak DV i MPEG, o których więcej w dalszej części tego rozdziału w podrozdziale „Kompresja stratna i bezstratna”). Rysunek 1.8 przedstawia przykład obrazu złożonego z siatki makrobloków.

Sytuacja przedstawia się następująco: kodek podzielił obraz na grupy bloków, uwzględniając kolejne klatki (pamiętaj, że wszystko to odbywa się w sposób dynamiczny). Wewnątrz bloków kodek ma więc informacje o poprzednich klatkach i dlatego też, aby zrekonstruować całą klatkę, musi jedynie zapamiętywać różnice pomiędzy tymi blokami, a nie całymi obrazami. Działa to bardzo efektywnie w przypadku wideo, które w ogóle nie zawiera scen dynamicznych lub ma ich niewiele. Przykładami takich nagrań wideo mogą być wywiady lub statyczne teksty.

Mimo złożoności całego procesu jest on stabilny i pozwala na uzyskanie bardzo dobrych rezultatów. Nie pozwala jednak na osiągnięcie tak dobrego współczynnika kompresji, aby zmniejszyć wielkość pliku wideo wysokiej rozdzielczo-

Rysunek 1.8.

Dzięki makroblokom kodeki mogą podzielić obraz na grupy pikseli, które są umieszczone obok siebie. Pozwala to na ich efektywniejsze przetwarzanie. Niektóre kodeki potrafią dzielić makrobloki na mniejsze grupy zwane partycjami



ści do znośnego już rozmiaru. W tym momencie przydają się podróże w czasie. Dowiedziałeś się już, że kompresja wideo poszukuje elementów wewnątrz klatki, które mogą być opisane w prosty sposób. Aby wygenerować cały obraz, wystarczy czasem niewiele informacji, by zapamiętać różnicę między kolejnymi klatkami. Jednakże w przypadku wideo zawierającego dużo ruchu, takiego jak ujęcia wykonywane za pomocą ręcznej kamery filmowej lub klip pokazujący eksplozję, pomiędzy ramkami następuje zmiana wielu pikseli (jak pokazano na rysunku 1.9), dlatego też więcej danych musi zostać przesłanych.

Rysunek 1.9.

Górna sekwencja klatek, zawierająca mnóstwo ruchu i niewiele powtarzających się elementów, jest trudniejsza do skompresowania niż sekwencja dolna, w której zmienia się stosunkowo niewiele



W górnym zestawie klatek futbolista porusza się po ekranie, a kamera śledzi jego ruchy. Oznacza to, że praktycznie każdy piksel ulega zmianie pomiędzy kolejnymi klatkami, dlatego też wykonanie wysokiej jakości kompresji bez użycia dużej ilości danych jest trudnym zadaniem. Z drugiej strony, sekwencja dolna skompresuje się dużo bardziej efektywnie. Kamera jest nieruchoma, a oprócz przedstawionego skoku motocyklisty inne elementy obrazu nie zmieniają się.

W taki właśnie sposób, ogólnie rzecz biorąc, działa kompresja. Istnieje oczywiście cały zakres bardziej specjalizowanej terminologii, która używana jest w celu opisanego, jak i co dzieje się podczas procesu kompresji. Ważne jest, aby zrozumieć pewne definicje przedstawione w następnych podrozdziałach, a także skutki zastosowania określanych przez nie działań w przypadku Twojej pracy, rzeczywistej

kompresji, w przesyłaniu informacji lub w procesie archiwizacji, który mógłbyś wykonywać dla swoich danych wideo. Nie musisz koniecznie uczyć się na pamięć urywków z tej książki, lecz powinieneś wiedzieć, co oznaczają dane terminy, gdy o nich usłyszysz. Im więcej będziesz się zajmować kompresją, tym częściej będziesz się z nimi spotykać w trakcie swojej codziennej pracy. Prawidłowe rozpoznawanie zwrotów może Ci jedynie pomóc.

Kompresja stratna i bezstratna

Wszystkie kodeki, które omawiam w tej książce (a także te, o których nie wspominałem), są bezstratne lub stratne. Podobnie jak ma to miejsce w przypadku dźwięku, **bezstratne** kodeki wideo potrafią po dekompresji danych uzyskać idealną co do każdego bitu kopię oryginału. Mimo że bezstratna kompresja wideo jest możliwa, nie jest jednak często stosowana. Wynika to stąd, że system kompresji bezstratnej może czasami wygenerować pliki (lub jego fragmenty), które są tak duże lub mają taką samą *prędkość przesyłu danych* (ang. *data rate*) jak nieskompresowany oryginał. W rezultacie cały sprzęt, używany w systemie bezstratnym, musi działać tak szybko, by obsłużyć wideo nieskompresowane, co od razu eliminuje wszystkie korzyści ze stosowania kompresji. Na przykład, cyfrowa taśma wideo nie może w prosty sposób zmieniać swojej prędkości przesyłu danych, dlatego praca z krótkimi pakietami wideo, transmitowanymi momentami z maksymalną szybkością, może być bardziej skomplikowana niż obsługa systemu o również maksymalnej, lecz niezmienną się wartości tejże prędkości.

Z drugiej strony, dane uzyskane po kompresji **stratnej** i poddane następnie dekompresji mogą się znacząco różnić od oryginału, lecz są mimo wszystko na tyle do niego podobne, że ich użyteczność jest istotna. Kompresja stratna jest najczęściej używanym sposobem kompresowania wideo, szczególnie podczas przygotowywania danych dla finalnego sposobu dostarczania, takiego jak DVD, Internet czy też urządzenia przenośne.

Co należy zapamiętać o kompresji stratnej i bezstratnej?

Gdy pracujesz ze swoimi projektami wideo, pamiętaj po prostu, że dane mogą być przechowywane w postaci stratnej i bezstratnej. Tryb bezstratny może, lecz nie musi, być przez Ciebie wykorzystany. Jeśli go wybierzesz, będziesz potrzebował dość zaawansowanego (i kosztownego) sprzętu, aby przechować w nim dane wideo, a następnie odpowiednio szybko je odtworzyć. Nawet jeśli użyjesz kompresji stratnej, możesz osiągnąć wyższą jakość (i większą prędkość przesyłu danych), niż zakładałeś na początku. Wideo, dostarczane w postaci końcowej, będzie prawdopodobnie miało format stratny. Pamiętaj, że nie oznacza to pogorszenia jakości, a jedynie to, że istnieje mniej danych składających się na obraz wideo.

Kompresja przestrzenna (DCT) i kompresja wavelet

Przedstawiłem już różnicę między kompresją stratną a bezstratną, a teraz zamierzam opisać dwa inne rodzaje kodeków: przestrzenne i oparte na analizie falkowej (wavelet). Kompresja przestrzenna (ang. *spatial compression*) jest podstawą tego, co poprzednio opisałem jako ogólną zasadę działania kompresji wideo — usuwaniem nadmiarowego materiału z kolejnych klatek z uwzględnieniem upływu czasu. Dyskretna transformata kosinusowa (DCT) jest odmianą kompresji przestrzennej. Jest ona zawsze stratna.

Kompresja wideo oparta na analizie falkowej (inaczej *kompresja wavelet*, ang. *wavelet compression*) jest nowoczesną odmianą kompresji obrazów i danych wideo, lecz nie jest zbyt często używana. Kompresja wavelet może być idealnie bezstratna lub stratna, tak jak DCT. Kodeki oparte na kompresji wavelet są rzadziej spotykane niż ich odpowiedniki używające kompresji DCT; utworzone przez nie pliki są często dużo większe i zwykle zużywają więcej mocy procesora podczas odtwarzania. Są one więc mniej odpowiednie do użycia w popularnych aplikacjach i platformach wideo. Jednym z przykładów kodeka wideo, opartego na analizie falkowej, jest JPEG 2000 uważany za bardzo dobry format pośredni.

Co powoduje, że JPEG 2000 jest lepszy od kodeków DCT? Trzy właściwości:

- *Bardzo dobra sprawność kompresji:* w przypadku wysokich prędkości transmisji, podczas których artefakty są prawie niezauważalne, obrazy generowane przez JPEG 2000 mają niewiele lepszą jakość w porównaniu z rezultatami działania innych kodeków. Dla niższych prędkości transmisji standard JPEG 2000 umożliwia jednak uzyskanie dużo lepszych wyników.
- *Kompresja stratna i bezstratna:* w przeciwieństwie do kodeków opartych na dyskretniej transformacie kosinusowej, kodeki wavelet mogą być zarówno stratne, jak i bezstratne, pozwalając dzięki temu na stosowanie szerokiej gamy opcji podczas pracy z plikami wideo.
- *Informacja przestrzenna w dodatkowych kanałach:* kodeki wavelet w pełni wspierają przezroczystość i kanały alfa.

Więcej przykładów dotyczących kodeków wavelet zobaczymy w rozdziale 3.

Co należy zapamiętać o kompresji przestrzennej i wavelet?

Kompresja przestrzenna, a szczególnie kompresja DCT, jest bardziej rozpowszechniona od kompresji wavelet, mimo że ta ostatnia zdobywa coraz większą popularność. W większości przypadków nie będziesz musiał się interesować tym, czy w pracy używasz tego czy innego rodzaju kodeka. Ze względu jednak na to, że standard kompresji przestrzennej jest obecnie wszechobecny, musisz zapewniać dostarczenie końcowej zawartości zakodowanej za jego pomocą.

Kwantyzacja

Kodeki oparte na dyskretnej transformacie kosinusowej (DCT) oraz na analizie falkowej (wavelet) wykorzystują kwantyzację. **Kwantyzacja** jest procesem zamieniającym ciągły zakres wartości na niewielki zbiór dyskretnych symboli lub wartości całkowitych; inaczej mówiąc, pozwala ona za pomocą matematycznie zdefiniowanego sposobu znaleźć efektywną metodę opisaną wszystkich pikseli na obrazie.

Kwantyzacja odgrywa znaczącą rolę w kompresji stratnej i może być traktowana jako podstawowy element, który odróżnia ją od kompresji bezstratnej.

Co należy zapamiętać o kwantyzacji?

W przypadku wideo kwantyzacja służy do ustalania poziomu jakości. Im mniejsza wartość kwantyzacji, tym lepsza jest jakość obrazu.

Kompresja międzyklatkowa i wewnątrzklatkowa

W tym rozdziale napisałem wcześniej, że kompresja potrafi śledzić zmiany pikseli między kolejnymi klatkami wideo, uwzględniając upływ czasu. Rodzaje kompresji międzyklatkowa i wewnątrzklatkowa umożliwiają zdefiniowanie metody pozwalającej określić zależności między różnymi klatkami.

Najbardziej powszechnym sposobem działania **kompresji międzyklatkowej** (ang. *interframe compression*) jest porównywanie każdej klatki wideo z poprzednią. Gdy oryginalna klatka wideo zostaje wycięta lub utracona (np. podczas edycji lub transmisji), ten rodzaj kompresji kopiuje dane z poprzedniej klatki, co powoduje, że kolejne obrazy nie mogą zostać poprawnie zrekonstruowane. Tylko wideo już przygotowane do dystrybucji (czyli takie, które nie może już być edytowane lub w jakikolwiek inny sposób modyfikowane) powinno być kompresowane do formatu międzyklatkowego.

Niektóre formaty wideo, takie jak DV, kompresują jednak każdą ramkę zupełnie niezależnie, używając w tym celu **kompresji wewnątrzklatkowej** (ang. *intraframe compression*). Edycja wideo skompresowanego w tym formacie jest podobna do edycji wideo nieskompresowanego, podczas której system przetwarzania danych nie musi wykorzystywać innych klatek, aby zdekodować informacje niezbędne do wygenerowania potrzebnego obrazu.

Inną różnicą między kompresją międzyklatkową a wewnątrzklatkową jest to, że w przypadku systemów wewnątrzklatkowych każda klatka używa w przybliżeniu takich samych ilości danych. W większości systemów międzyklatkowych niektóre klatki, zwane **I-klatkami** (ang. *I-frame*), nie przenoszą informacji z innych obrazów i dlatego też wymagają dużo więcej danych do utworzenia niż inne klatki znajdujące się obok. Rozdział 2. przedstawia więcej informacji na temat istniejących rodzajów klatek, włączając w to I-klatki.

W przypadku edytorów nielinowych, takich jak Apple Final Cut Pro i Adobe Premiere Pro, można obecnie zidentyfikować problemy występujące podczas braku wymaganych I-klatek. Pozwoliło to nowszym formatom, takim jak wideo wysokiej rozdzielczości (HDV) wykorzystującemu format MPEG-2, na edycję bez potrzeby wcześniejszej konwersji na pośrednie formaty I-klatkowe. Jednakże proces ten wymaga większej mocy obliczeniowej niż edycja wideo wewnątrzklatkowego o tej samej jakości obrazu, gdyż procesor w sposób ciągły przeszukuje inne klatki, aby odnaleźć informację niezbędną do odtworzenia tego obrazu, który ma być edytowany przez użytkownika.

Co należy zapamiętać o kompresji międzyklatkowej i wewnątrzklatkowej?

Kompresja międzyklatkowa działa poprawnie w przypadku wideo, które ma być po prostu odtwarzane przez użytkownika, lecz może spowodować problemy podczas próby edycji danych. Wideo, które jest wciąż w fazie produkcji i dlatego też może w każdej chwili zostać zmodyfikowane, powinno być kompresowane za pomocą metody wewnątrzklatkowej (takiej jak MPEG-2 lub DV).

Kompresja dźwięku

W czasie tej całej analizy dotyczącej kompresji bardzo łatwo można zapomnieć, że wideo ma najczęściej również dźwięk. W rzeczywistości wielu producentów filmowych powie Ci, że podczas oglądania filmu dźwięk jest ważniejszym elementem od obrazu. Na pewno nie jest on mniej ważnym czynnikiem, mimo że proces nagrywania i kompresji dźwięku nie jest tak skomplikowany jak w przypadku wideo. Wynika to jednak z tego, że ścieżki audio mają mniej danych do obróbki w porównaniu ze ścieżkami wideo.

Zwrot **kompresja** ma kilka znaczeń w technice audio, dlatego dobrze jest o tym pamiętać, szczególnie gdy rozmawiamy ze specjalistą z tej dziedziny. Istnieje stratna kompresja danych (podobna do tej, którą omówiłem w przypadku wideo), kompresja poziomu dźwięku oraz ograniczanie dynamiki dźwięku. Na temat kompresji oraz ograniczania dynamiki powiem więcej w rozdziale 4., omawiając techniki przetwarzania wstępnego dla dźwięku.

Podobnie jak w przypadku kompresji wideo, do kompresji audio również stosowane są algorytmy stratne i bezstratne. Kompresja stratna używana jest do codziennych zastosowań (takich jak odtwarzanie MP3). W obu tych rodzajach algorytmów następuje odpowiednia redukcja informacji nadmiarowej, aby zmniejszyć jej ilość potrzebną do opisanie danych. W tym celu używa się takich metod jak kodowanie, rozpoznawanie wzorców i prognozowanie liniowe.

Kompresja audio może jednakże szybko dać znać o sobie, a rezultaty mogą być całkiem zauważalne, szczególnie w muzyce, która ma więcej dynamicznych dźwięków niż słowo mówione. Muzyka jest podobna do wideo z dużą zawartością ruchu; biorąc pod uwagę bardziej dynamiczną akcję, kompresory muszą zajmować się większą ilością informacji w danym momencie. W przypadku monologu, w którym dzieje się dużo mniej, wymagane jest również mniej danych do jego nagrania. Gdy audio o wysokim stopniu dynamiki zostanie zapisane w zbyt małej liczbie bitów, końcowym rezultatem może być zniekształcony dźwięk — metaliczny, mający echo lub niezawierający basów.

Jak wcześniej wspomniano, dobrze wykonana kompresja audio jest tak samo ważna jak kompresja wideo, a przy wzrastającej popularności podkastingów wielu ludzi zacznie zwracać na nią większą uwagę, niż miało to miejsce w przeszłości. Kluczem do uzyskania dobrej ścieżki dźwiękowej, być może ważniejszym niż w przypadku wideo, jest uzyskanie dobrego materiału źródłowego.

Szacowanie wideo dla celów kompresji

Pierwszą część tego rozdziału poświęciliśmy na zapoznanie się z teoretyczną stroną systemów wideo oraz kompresji wideo. Mając tę wiedzę, w tym podrozdziale będziesz mógł połączyć teorię z praktyką i oszacować wideo, nad którym pracujesz, aby dość dokładnie określić, jak dobrze może zostać ono skompresowane. Następnie, w dalszej części książki, poznasz sposoby pozwalające na wykonanie samej kompresji.

Brzmi to dość zabawnie, ale będziesz musiał dosłownie usiąść i obejrzeć materiał wideo, który zamierzasz kompresować. Czasem pomijałem tę czynność — bywałem tak zajęty pracą, że gdy otrzymywałem taśmę do zakodowania lub klip w programie Final Cut Pro, który musiał być wyeksportowany, natychmiast zajmowałem się opcjami i wszystkimi tymi „matematycznymi” ustawieniami, nie interesując się w ogóle, jak wygląda samo wideo.

Związane z jakością wideo pojedyncze kwestie, które możesz ocenić jedynie przez zapoznanie się z nimi za pomocą własnych oczu, jak również praktyczna znajomość materiału źródłowego są często tak ważne jak posługiwanie się samą technologią wideo. Zanim zajmiesz się konkretnymi opcjami, dotyczącymi określonej techniki kompresji, wiedza na temat tego, co zamierzasz kompresować, jak Twoje wideo powinno wyglądać i co zamierzasz z nim dalej zrobić po wykonaniu operacji kodowania, pomoże Ci później w wyborze odpowiednich ustawień oraz filtrów, których będziesz chciał użyć, by uzyskać najlepszą jakość obrazu przy zachowaniu optymalnej wielkości pliku. Ogólnie rzecz biorąc, istnieją trzy podstawowe pytania, które zadaję sobie i innym podczas kompresji wideo. Jeśli potrafię na nie odpowiedzieć, mogę również wygenerować plik o lepszych parametrach.

W jaki sposób wideo zostało zapisane?

Decyzja dotycząca sposobu kompresji musi być jakoś związana ze sposobem, w jaki dane wideo zostało nagrane. Czynniki, które przyczynią się do wyboru rodzaju kompresji, mogą być rozłożone na dwie główne kategorie: sprzęt i produkcję. Sprawy związane ze sprzętem zawierają w sobie pytania dotyczące rodzajów i jakości kamer służących do nagrania wideo — im sprzęt był bardziej profesjonalny, tym lepsze materiały źródłowe najczęściej otrzymasz.

Do zapamiętania: kwestie dotyczące sprzętu

Czy wideo zostało nagrane za pomocą sprzętu profesjonalnego, czy też zwykłej kamery?

W jakim formacie nagrano wideo?

Czy wideo zawiera ścieżkę dźwiękową?

Jaki był rodzaj oświetlenia: profesjonalne lub naturalne?

Do zapamiętania: kwestie dotyczące produkcji

Czy obrazy wideo zawierają ruch, czy też są statyczne?

Czy zdjęcia były wykonywane metodą z przeplotem, czy progresywną?

Czy ruchy kamery były płynne, czy gwałtowne?

Czy obrazy wideo mają wysoki, czy niski kontrast?

Czy nagrania zawierają w większości obiekty odległe, czy raczej bliskie?

W jaki sposób przeprowadzono edycję wideo?

Gdy wideo zostało już nagrane, należy zadać pytanie, w jaki sposób było ono przetwarzane na etapie postprodukcji. Sposoby edycji są bardzo różne w zależności od zawartości materiału źródłowego, wymaganego produktu końcowego oraz konkretnego edytującego specjalisty. Pewne rodzaje edycji prowadzą w sposób naturalny do takiego typu wideo, który może zostać łatwiej skompresowany niż inne. Jeśli będziesz potrafił rozróżnić rodzaje edycji, pozwoli Ci to na poświęcenie czasu właściwym plikom (lub nawet zmianę sposobu edycji, jeśli to będzie możliwe).

Do zapamiętania: kwestie dotyczące postprodukcji

Czy wideo zostało wydłużone, czy skrócone?

Czy wideo zawiera cięcia lub przejścia (np. efekty przenikania)?

Czy wideo zawiera dużo elementów przeźroczyстых lub warstw?

Czy obrazy wideo mają wysoki, czy niski kontrast?

Czy jakość wideo została świadomie obniżona za pomocą odpowiedniej techniki lub też czy zostały zmodyfikowane kolory (jest to częste zjawisko podczas produkcji teledysków muzycznych i może Cię ono bardzo zirytować, jeśli nie jesteś jego świadomy)?

Kto ma być docelowym odbiorcą?

Tak więc Twoje wideo zostało zapisane, poddane edycji i jest gotowe do rozpowszechnienia. W tym momencie należy rozpatrzyć jeszcze jedną sprawę — kto będzie oglądać Twoje wideo i gdzie będzie się to odbywać? Mnóstwo programów wideo odbieranych jest za pomocą telewizji, lecz jeśli zajmujesz się kompresją, robisz to, aby rozpowszechnić dane przy użyciu także innego medium (lub najczęściej wielu mediów).

Do zapamiętania: kwestie dotyczące dystrybucji

W jaki sposób widzowie będą oglądać wideo: na komputerze, w telewizji, w kiosku wideo lub na urządzeniu przenośnym?

Czy będzie to pokaz dla pojedynczego widza, czy dla grupy widzów?

Czy dźwięk wideo będzie mógł być łatwo odbierany (inaczej mówiąc, czy miejsce odtwarzania wideo jest spokojnie, czy hałaśliwe)?

Jak duży jest ekran i w jakiej odległości od niego będą znajdować się widzowie?

Co jest ważniejsze: jakość czy prędkość transmisji?

Zakończenie

Mówiąc ogólnie, specjaliści kompresji nie mogą być absolutystami. Nie istnieje żadna prosta odpowiedź (lub plan działania) na wiele zadawanych pytań. Zamiast tego należy postrzegać kompresję jako działanie polegające na analizowaniu informacji oraz utrzymywaniu równowagi pomiędzy wymaganiami dotyczącymi jakości obrazów a rzeczywistą metodą dostarczania wideo.

Będąc świadomym istnienia elementów związanych z produkcją wideo oraz — co ważniejsze — tego, w jaki sposób wpływają one na rezultaty Twojej pracy, możesz zwiększyć prawdopodobieństwo wykonania poprawnej kompresji wideo. Będziesz znał pułapki i zagrożenia czekające na Ciebie i dzięki temu przygotujesz się lepiej do analizy wyników, które możesz osiągnąć po przeprowadzeniu kompresji.

Formułowanie pytań przed zaprojektowaniem i implementacją opcji kompresji może pomóc Ci wykonać wideo wysokiej jakości podczas przeprowadzania końcowych operacji.

Jak powiedziałem, kompresja wideo to coś znacznie więcej niż tylko sam opis jej działania. Następnym krokiem będzie zrozumienie, w jaki sposób język kompresji może wpłynąć na sposób tworzenia Twojego wideo oraz jak należy przeprowadzać jego kodowanie.