
Mieszana Zawartość Rastrowa

Praktyczne zastosowanie modelu kompresji w formacie DjVu

Skanowanie dokumentu

- Podczas skanowania stron dokumentu standardem są ustalone parametry DPI i system koloru
- Z reguły dokument skanuje się w rozdzielczości nie mniejszej niż 300DPI
- Dokument można skanować w czerni i bieli, odcieniach szarości i kolorze
- Przy czarnobiałym obrazie DPI powinno być większe (np. 600DPI)
- Kolorowe strony są skanowane do systemu kolorów RGB

Kolor dokumentu

- Obraz w czerni i bieli zawsze zajmuje mało miejsca przestrzeni dyskowej
- Obraz czarnobiały wymaga jednego bita na piksel
- Obraz w RGB wymaga 24 bitów na piksel (bez kanału alfa)
- Niektóre zasoby nie wymagają uwzględnienia koloru przy udostępnianiu zbiorów, mimo to skany źródłowe (tzw. mastery) powinny być wykonane w kolorze
- Przy udostępnianiu zbiorów najważniejsze jest przekazanie treści udostępnianego dokumentu
- Kolor i wygląd dokumentu również mogą być treścią
- Co może być treścią dokumentu?

Treść dokumentu

- Tekst
- Karta (m.in. jej wymiary)
- Obraz
- Kolor tekstu
- Kolor obrazu
- Układ dokumentu
- Cechy fizyczne nośnika treści (np. materiał tworzący kartę)
- Obiekty dołączone do dokumentu (sfragistyka)

Sposoby zmniejszenia rozmiaru

- Kompresja w dokumencie może być bezstratna, wtedy zastosowanie takiego algorytmu nie ma konsekwencji dla obrazu
- Stratna kompresja zmienia sam obraz, zniekształcając go i upraszczając
- Kompresję stratną można stosować dopiero na końcu procesu obróbki graficznej
- Grafika rastrowa ma dwa rodzaje algorytmów kompresji zależne od systemu kolorów w obrazie
- Osobne są algorytmy dla plików kolorowych i dla czarnobiałych
- W jednym pliku mogą być stosowane dwa rodzaje kompresji

Sposoby kompresji - bezstratne

- Kompresja poprzez korzystanie z ustalonych deskryptorów (ZIP)
- Zliczanie powtórzeń (RLE)
- Tworzenie słownika dla pojedynczego pliku (LZ77)
- Tworzenie słownika dla pliku, którego dekodowanie odbywa się na podstawie heurystyki kodu tego pliku (LZ78)
- Tworzenie słownika poprzez konkatencje wyrażen (LZW)
- Tworzenie konkatencji wyrażen słownika w jednym wpisie w indeksie (LZMW)
- Tworzenie słownika konkatencji wraz z modyfikowalnymi prefiksami (LZAP, szybko rozrastający się słownik)
- Kompresja z zastosowaniem predykcji ciągu wyrażen (LZP)
- Parowanie znaków z buforu (LZSS)
- Predykcja częściowego dopasowania (PPM)

Transformaty, algorytmy

- Inną metodą jest tworzenie indeksu danych (np. niesłownikowych) i wektoryzacja ciągu znaków słownikowych w indeksie

Transformata Burrowsa-Wheelera

- Czasem w kodowaniu wykorzystuje się więcej niż jeden algorytm (np. JPEG wykorzystuje m.in. Kodowanie Huffmana)
- Połączenie wielu algorytmów rzadko przynosi dobre rezultaty

Cechy kompresji bezstratnej

- Kompresja bezstratna jest uniwersalna, tzn. algorytmy można zastosować w każdym rodzaju pliku
- Kompresja bezstratna zapewnia dokładność zapisu danych
- Wykorzystanie zasobów sprzętowych, wielowarstwowość algorytmów i dostęp do danych jest taki sam, jak w przypadku kompresji stratnej

Sposoby kompresji stratnej I

- Zmiana kolorów
- Ludzkie oko zauważa różnice pomiędzy kolorami skupiając się bardziej na iluminacji poszczególnych sektorów niż oceniając rzeczywiste odchylenia barwy
- Zmianie kolorów podlega więc system RGB, w którym zamienia się kodowanie 24 bitowe na inne
- Taka zmiana jest równa zmniejszeniu ilości barw wyświetlanych przez monitor
- Kompresji podlega zmniejszenie przedziału wartości dla kanału

Sposoby kompresji stratnej II

- Drugi poziom to stosowanie kompresji DCT (dyskretna transformacja kosinusowa)
- Metoda polega na podziale danych na zwarte bloki (w obrazie np. 8x8, 16x16, 32x32 pikseli)
- Pojedyncze bloki są uśredniane, a wartości średnie dla bloku są zapisywane tylko jeden raz
- Dalej każdy piksel jest jedynie odchyleniem wartości uśrednionej
- Metoda ta sprawdza się głównie w grafice rastrowej (także wideo)
- Następnym poziomem jest zastępowanie wartości pojedynczego bloku różnicą pomiędzy następującymi blokami

Sposoby kompresji stratnej III

- Kompresja stratna bazuje na kwantyzacji danych
- W przypadku wartości średnich dla bloku często są one zapisane jako *loop/float*
- Wówczas kompresja nie daje efektów w postaci zmniejszenia ilości zajmowanej przestrzeni dyskowej
- Wszystkie ostateczne dane są zaokrąglane do liczb rzeczywistych (*integer*)
- Zaokrąglanie jest powodem największych strat jakości obrazu
- Aby zminimalizować efekt przejścia blokowego pobiera się część próbek wspólnie dla wielu bloków (tzw. MDCT, nie używane w kompresji JPG)

Sposoby kompresji stratnej IV

- Przy kompresji korzystającej z DCT niewielkiej wielkości bloki dają efekt agresywnych przejść na granicy dwóch bloków
- Aby złagodzić ten efekt dobrze jest powiększyć blok, wówczas wartości średnie pomiędzy blokami będą zbliżone
- Wartości skrajnie różne są kodowane za pomocą Kodowania Huffmana (czyli bezstratnie)

Sposoby kompresji stratnej V

- Inną metodą jest kompresja poprzez falki (ang. Wavelet, DWT)
- Polega ona na stosowaniu funkcji macierzystej z przesunięciem dla poszczególnych zbiorów wartości (zbiór danych jest zawsze odniesieniem do poprzedniego i do funkcji macierzystej)
- Kolejnym poziomem kompresji jest zastosowanie Transformaty Fouriera (tak jak przy kompresji bezstratnej Transformaty Burrowsa-Wheelera)
- Falki kompresując obraz nie dają efektu przejścia blokowego

Sposoby kompresji stratnej VI

- W kompresji obrazu stosowanie kompresji wszystkich kanałów jednocześnie daje lepsze efekty
- JPEG2000 – dzielenie obrazu na warstwy różnych częstotliwości
- Każda z warstw podlega Transformacji Falkowej, dzięki czemu różnice pomiędzy wartościami są mniejsze
- Mniejsze różnice pozwalają na dokładniejsze odwzorowanie obrazu
- Mniejsze różnice powodują znaczne uproszczenie transformaty i ułatwiają zliczanie powtórzeń

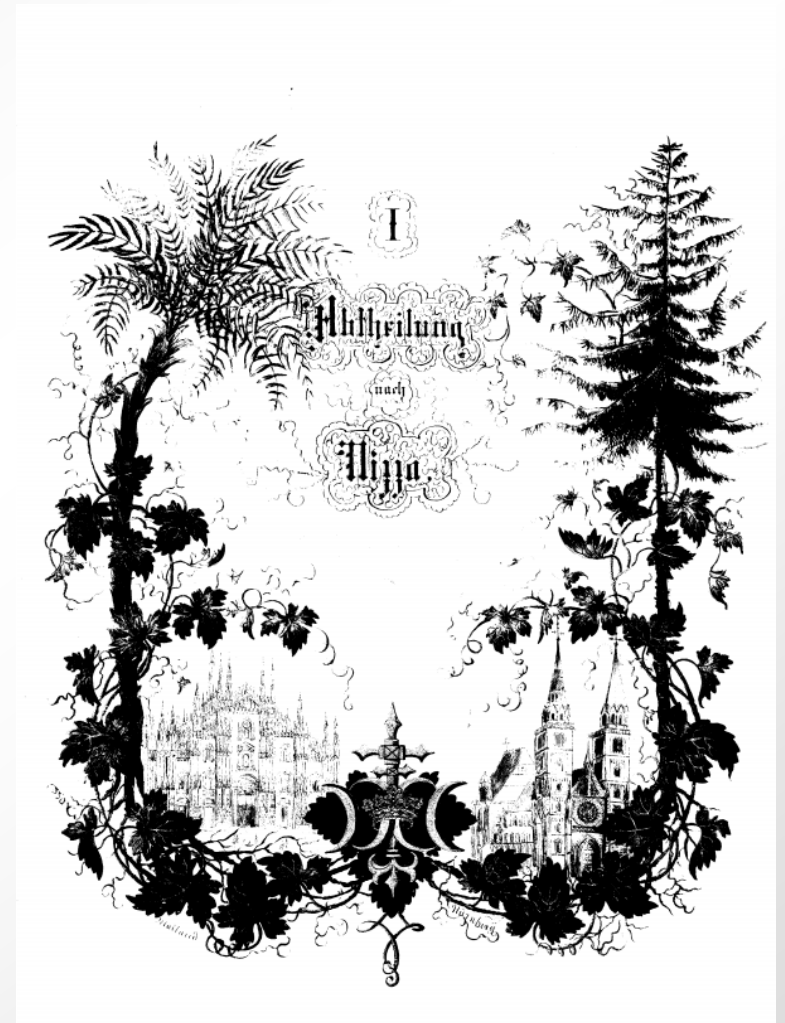
Kompresja dwupoziomowych danych

- Obraz czarnobiały ma jedynie zapis binarny
- Zapis binarny ogranicza się do kodowania arytmetycznego
- Kodowanie arytmetyczne polega na entropii (średnia wartość z przedziału jest podstawą kodowania)
- Używa się dwóch algorytmów kompresji binarnej, JBIG (obecnie JBIG2) oraz CCITT (v4, v6)
- Koder JBIG2 pozwala na konwertowanie obrazów z koloru do czerni i bieli w sposób kompleksowy
- Obraz kolorowy może być zapisany na różne sposoby

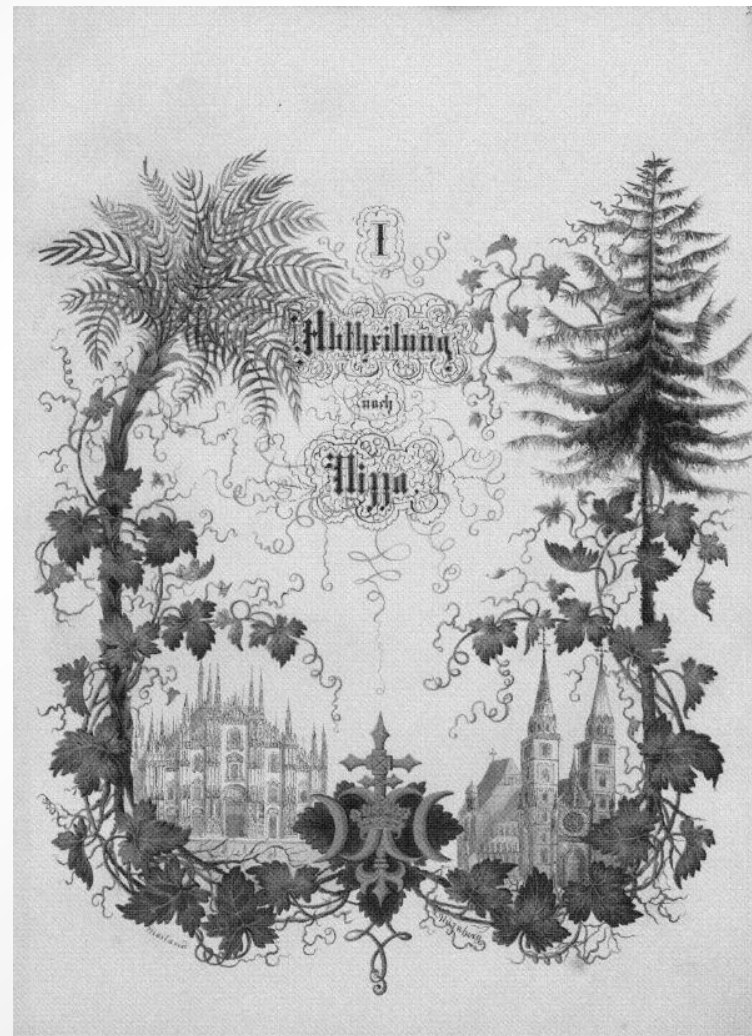
Rezygnacja z koloru

Konwersja do czerni i bieli

Zwykły



Z półtonami



Z półtonami



Kompresja dokumentu

- Obrazy dokumentów z reguły mają wiele cech wspólnych
- Dzięki temu można stworzyć model kompresji odpowiadający specyficznym potrzebom dokumentu tekstowego
- W dokumencie tekst jest najważniejszy, pozostałe elementy z reguły uzupełniają treść
- W związku z tym najważniejsze jest odtworzenie liter w dobrej jakości, pozostałe elementy nie wymagają tak dobrej jakości (np. kolor czcionki z reguły jest jednolity)
- Najlepszą kompresją byłoby zapisanie samego tekstu bez obrazów (wymaga to jednak dokładnego rozpoznania OCR)

Dokument czarnobiały

ELECTRUM • Vol. 4
Kraków 2000

Tadeusz Kotula

PRINCIPALES CIVITATUM:
GRUPA SPOŁECZNA I MUNICYPALNA INSTYTUCJA
EPOKI PÓŹNEGO CESARSTWA RZYMSKIEGO

W piątej księdze traktatu *De cognitionibus*, fragmentarycznie zachowanego w Justynianowych *Digestach*, pisał Kallistrat, jurysta okresu Sewerów, o *decuriones et principales civitatum, qui capitale admiserunt*¹. Mamy tu niewątpliwie do czynienia, już na przełomie II/III wieku, z oficjalną instytucją cesarstwa. Zaczniemy od krótkiego stwierdzenia, że termin *principales* ma sens *par excellence* hierarchiczny – na określenie grupy społecznej, wysuwającej się, jak sama nazwa wskazuje, na pierwszy plan.

Spółczesność rzymskie nigdy nie było demokratyczne, także pod względem strukturalnym. Już w dobie republiki wyodrębniły się z rządzącej *nobilitas* wielkie rody o przewodniej roli w życiu politycznym. Poświęcił im zasłużoną uwagę niemiecki historyk F. Münzer i Włoch F. Cassola. Wśród samej arystokracji senatorskiej oligarchiczne rodziny odcięły się od *senatores pedanei*. W epoce cesarstwa proces społecznego rozwarstwienia znacznie się pogłębił, o czym świadczyło mnożenie się na samym szczycie społecznej hierarchii wszelkich obojętności i wariantów liczebniaka porządkowego *primus: primarii, primates, primores, principes, principales, priores* oraz ich odpowiedników w ogólnikowym znaczeniu czołowych osobistości, *civitatit vertices*. Jak wszystkie tego rodzaju określenia z wokabularza społecznej stratyfikacji, także pojęcie *principales* nie miało zrazu znaczenia terminu technicznego w sensie instytucjonalnym. W dobie Antoninów określił Apulejusz kartagińskich dekurionów mianem *principes Africae viri*, a podobny ogólny zwrot odniósł do najznamienitszych przedstawicieli arystokracji municypalnej (*principales viri*), zaś Tertulian wyróżniał wśród kartagińskich chrześcijan *principales personas*².

Już jednak za Sewerów, jak wyżej wspomniano, prawnicy rzymscy uważali kategorię miejskich *principales* wydzieloną z ogółu dekurionów za formalną instytucję. Zacytujmy szerzej odnośny tekst Kallistrata: *De decurionibus et principalibus civitatum, qui capitale admiserunt...* (i dalej): *Si qui ex principalibus alicuius civitatis latrocinium fecerint...* Surowe wyroki spadały zatem również na najwyższych miejskich notabłów, którzy popełnili cięż-

¹ Dig. XLVIII, 19, 27, 1–2.

² Apul. Flor. 16, 72; 21, 100; por. Metam. IV, 26: *speciosus adulescens, inter suos principalis*; Tert. ad Scap. 5.

ELECTRUM • Vol. 4
Kraków 2000

Tadeusz Kotula

PRINCIPALES CIVITATUM: Grupa społeczna i municypalna instytucja
EPOKI PÓŹNEGO CESARSTWA RZYMSKIEGO

W piątej księdze traktatu *De cognitionibus*, fragmentarycznie zachowanego w Justynianowych *Digestach*, pisał Kallistrat, jurysta okresu Sewerów, o *decuriones et principales civitatum, qui capitale admiserunt*¹. Mamy tu niewątpliwie do czynienia, już na przełomie II/III wieku, z oficjalną instytucją cesarstwa. Zaczniemy od krótkiego stwierdzenia, że termin *principales* ma sens *par excellence* hierarchiczny – na określenie grupy społecznej, wysuwającej się, jak sama nazwa wskazuje, na pierwszy plan.

Spółczesność rzymskie nigdy nie było demokratyczne, także pod względem strukturalnym. Już w dobie republiki wyodrębniły się z rządzącej *nobilitas* wielkie rody o przewodniej roli w życiu politycznym. Poświęcił im zasłużoną uwagę niemiecki historyk F. Münzer i Włoch F. Cassola. Wśród samej arystokracji senatorskiej oligarchiczne rodziny odcięły się od *senatores pedanei*. W epoce cesarstwa proces społecznego rozwarstwienia znacznie się pogłębił, o czym świadczyło mnożenie się na samym szczycie społecznej hierarchii wszelkich obojętności i wariantów liczebniaka porządkowego *primus: primarii, primates, primores, principes, principales, priores* oraz ich odpowiedników w ogólnikowym znaczeniu czołowych osobistości, *civitatit vertices*. Jak wszystkie tego rodzaju określenia z wokabularza społecznej stratyfikacji, także pojęcie *principales* nie miało zrazu znaczenia terminu technicznego w sensie instytucjonalnym. W dobie Antoninów określił Apulejusz kartagińskich dekurionów mianem *principes Africae viri*, a podobny ogólny zwrot odniósł do najznamienitszych przedstawicieli arystokracji municypalnej (*principales viri*), zaś Tertulian wyróżniał wśród kartagińskich chrześcijan *principales personas*².

Już jednak za Sewerów, jak wyżej wspomniano, prawnicy rzymscy uważali kategorię miejskich *principales* wydzieloną z ogółu dekurionów za formalną instytucję. Zacytujmy szerzej odnośny tekst Kallistrata: *De decurionibus et principalibus civitatum, qui capitale admiserunt...* (i dalej): *Si qui ex principalibus alicuius civitatis latrocinium fecerint...* Surowe wyroki spadały zatem również na najwyższych miejskich notabłów, którzy popełnili cięż-

¹ Dig. XLVIII, 19, 27, 1–2.

² Apul. Flor. 16, 72; 21, 100; por. Metam. IV, 26: *speciosus adulescens, inter suos principalis*; Tert. ad Scap. 5.

PRONUMERATA WYNOŚI:

we Lwowie:
miesięcznie 2 korony, za dwurazową dost. do domu dopłaca się 60 h.

na prowincyi:

rocznik 26 k. 40 h. z 2-krótną wys. 32 k. — b.
kwartal. 6. 60 " " " 8 " —
miesięcz. 2. " " " 2. 70 "

W Niemczech: miesięcznie . 4 kor.
W innych krajach . 6 "

Pronumeratcy Słowa Polskiego
płać za Wiek XX. 1 kor.

Zaśada amlana adresu 40 h.

Słowo Polskie

wychodzi 2 razy dziennie.

OGŁOSZENIA:

Za 1 wiersz petiowy albo jego miejsce 20 h., nadstawia wiersz garmondem 80 h., małe ogłoszenia za wyraz 6 h., najmniej 60 h.

Numer pojedynczy:

We Lwowie 6 h.
na prowincyi 8 "

Adres: „Słowo Polskie“

Lwów, ul. Chorążczyzny 1. 17.

TELEFON 641.

Stalowa płyta.

Niezawodnie do interesujących postaci wśród czeskich polityków należy małutki, pękaty, czerwony, ruchliwy, o wiecznie zachrypłym głosie, wiceprezes czeskiego klubu — poseł Pacak. Niedługo, gdy nie konfiskowano gazet, ale wsadzano w Czechach redaktora do kozy, Pacak, który pisać nie umiał i nie pisał; był redaktorem od siedzenia w kozie i siadywał. Wyszedł potem na porządnego adwokata, robił czupurnie polityką i został posłem. Sława jego jednak, która doprowadziła małutkiego Pacaka na krzesło wiceprezesa czeskiego klubu — to — mała broszurka, napisana w kąpielach morskich, w której zebrał materiały do rozwiązania sprawy językowej i naszkicował rozwiązanie tak — jak się stało czynem w rozporządzeniu hr. Badeniego.

Ten małutki, dobroduszny niby, filuterny, łatwo zapalny, nie zgodliwy Pacak, narobił już dużo kłopotów w Austrii. A dzisiajse depresję z Pragi donoszą, że cisnął w czeskim Sejmie bombę, która ma rozzerwać wszystkie dotychczas nawiązane nitki ugodowe, naprzód zatargasować drogę akcji pacyfikacyjnej rządu i rozsądzić parlament.

Ten małutki Pacak należy do najzręczniejszych nieszczęśliwów czeskiego obozu. Towarzyszy mu opinia „jednego z najumiarkowańszych“, a życie jego idzie wśród ciągłych — powodzeń. Uosobienie dobrodusznosci i filuterni umie „brać“ ludzi na pierwszą, jak wziął hr. Badeniego, jak umiał brać Polaków. Metoda jego, to kombinacja piorunów i uśmiechów, chrzypłego ryku lwa (z tych mniejszych) i uścisków, przechodzi z jednego w drugie od razu — i ma samych przyjaciół. Dla Polaków on jest „umiarkowany moskalofil“, Wolf nazywa go „mein lieber Pacaker“, jest w stanie huczoć opozycyjną mowę przeciw ministrowi i w pół minuty potem ścisnąć go. Thuna, gdy był namiestnikiem Czech, porównywał z Hei-

broni się przed czechizacją. Nie wolno zapominać, że w tych centrach niemieckich — jest 75 powiatów sądowych, w których ludność czeska nie wynosi 3 procent, w 25 powiatach nie dochodzi jednego procentu.

W tych zamkniętych okręgach jest na 1,592 000 Niemców zaledwie 51.000-Czechów. Mimo niesłychanych rezultatów narodowego czeskiego odrodzenia, mimo imponującej żywotności, działalności, masowości, rozpiętej się po całym kraju i wyciuwającej rokrocznie ogromny nadmiar do prowincji niemieckich, dalsze zdobycze na terytorium niemieckim w Czechach, są tylko możliwe przy pomocy całej maszyny państwowej, gdyby państwo uznało za wskazane użyć całej swej siły dla przyspieszenia czechizacji niemieckich części Czech.

Od 1897 przechodzi państwo straszne przesilenie, Austria ginie, pokazało się niebezpieczeństwo, że Niemcy udusić się nie dadzą — trzeba było zawrócić z fałszywej drogi i zaczęło się zanosić na jakieś ugodowe załatwienie. Niemcy, obok rozmaitych przesadnych pretensyj, które dało się obciąć, sami ofiarują język czeski w wewnętrznej służbie w czeskich okręgach. Myśli językowego rozgraniczenia terytorjalnego zaczęła się przyjmować, świeżo nawet Kramarz przyjął podział na językowe strefy. Na to przychodzi znowu mały Pacak i ciska żągiel nowego pożaru.

Zdawałoby się więc, że nie „stalowa płyta“ ale przepaść nie do wypełnienia otwarła się znowu pomiędzy walczącymi obozami, a do przepaści ma runąć wszystko: konstytucja, parlamentaryzm, całość monarchii, Austria. Z niesłychaną lekkoomyślnością, butą i arogancją idą znowu Czesi na awanturę, zaślepieni jakąś mistyczną wiarą, że się skończy ostentacyjnym zwycięstwem, że wymuszają zdobycze, chociażby po trupie Austrii. Więc znowu z całą siłą grozi się pomocą rosyjską i wali cepem husyckim w ustrój państwa.

chunek i zważyć, czy żądane od nas ofiary stoja w jakimkolwiek stosunku do korzyści narodowych, politycznych, materialnych.

Przekroczenie tej granicy dla fałszywej polityki, dla jakichś widoków słowiańskich, którym się trzeba strasznie dokładnie przypatrzeć, byłoby ze szkodą Polaków, prostem szaleństwem i zbrodnią, popełnioną na swoich własnych narodowych interesach nie tylko w Austrii.

Nasze zaślepienie zdemoralizowało już Czechów w tym stopniu, że sobie z naszych zastrzeżeń nic nie robią. Jest to ostatni czas, żeby im powiedzieć: *sunt denique fines*.

Albo prawdą jest, że dla Polaków Austria ma jeszcze jakąś wartość, albo to już stanowisko przedawnione. Jeżeli ma wartość jeszcze, a interes narodowy chciałby tę wartość zachować na dalszą przyszłość, to nie wolno dłużej igrać z ogniem, bo nam panowie Pacaki spalą dom.

Włości rentowe.

Z kół poselskich otrzymujemy następujące uwagi:

Obrady Sejmu po feriach świątecznych rozpoczynają się pierwszeństwem czytaniem wniosku posła Potoczka o włościach rentowych.

Wniosek ten, jak wszystko, co dąży do poprawienia stosunków gospodarczych naszego włościanstwa, można tylko sympatycznie powitać — i nie przypuszczamy na chwilę, żeby Sejm chciał z nim tak postąpić, jak się to stało z wnioskiem posła Wójcicka, odrzuconym *a limine*, bo nie odesłanym nawet do komisji. Wniosek posła Potoczka tak, jak jest, nie może być przez Sejm przyjęty — ale są w nim zdrowe ziarna, które należy na polu naszych gospodarczo-społecznych stosunków, wypiełgnować i rozwinąć.

Rozpoznany tekst

lr, 191.

Lwdw. Środa, 25. kwietnia 1900.

Rok V.

PRENUMERATA WYNOŚI:

we Lwowie:
miesięcznie 2 korony, za dwurazową dost. do doWu dopłaca ale 60 h.

na prowincyi:

rocznie 26 K. 401, z 2-krotną wys. 321. — i.
kwartał 6 „ 60 „ . 8 „ —
miesięcz. 2 „ 2 „ 70 „

W Niemczech: miesięcznie . 4 kor.
W innych krajach „ . 6 „

Przenumeratory *Słowa Polskiego*
płacą sa *Wiek XX*, 1 kor.

Zażda. zmiana adresu 40 h.

sas*ś

Stalowa płyta.

Niezawodnie do interesujących postaci wśród czeskich polityków należy małutki, pękaty, czerwony, ruchliwy, o wieczne zachrypłym głósie, wiceprezes czeskiego klubu — poseł Pacak. Niegdyś, gdy nie konfiskowano gazet, ale wsadzano w Czechach redaktora do kozy, Pacak, który pisać nie umiał i nie pisał, był redaktorem od siedzenia w kozie i siedywał. Wyszedł potem na porządnego adwokata, robił czupumie politytę i został posłem. Sława jego jednak, która doprowadziła małutkiego Pacaka na krzesło wiceprezesa czeskiego klubu — to — mała broszurka, napisana w kąpielach morskich, w której zebrał materiał do rozwiązania sprawy językowej i naszkicował rozwiązanie tak — jak się stało czynem w rozporządzeniu hr. Badeniego.

Ten małutki, dobroduszny niby, filuterny, łatwo zapalny, ale zgodliwy Pacak, narobił już dużo kłopotów w Austrii. A dzisiejsze depesze z Pragi donoszą, że cisnął w czeskim Sejmie bombę, która ma rozarwać wszystkie dotychczas napisane niemi

Słowo Polskie

wychodzi 2 razy dziennie

broni się przed czechizacją. Nie wolno zapominać, że w tych centrach niemieckich — jest 75 powiatów sądowych, w których ludność czeska nie wynosi 3 procent, w 25 powiatach nie dochodzi jednego procentu.

W tych zamkniętych okręgach jest na 1,592 000 Niemców zaledwie 51.000-Czechów. Mimo niesłychanych rezultatów narodowego czeskiego odrodzenia, mimo imponującej żywotności, dzielności, ninożliwości, rozpiętej się po całym kraju i wylewającej rokrocznie ogromny nadmiar do prowincyj niemieckich, dalsze zdobycze na terytorium niemieckim w Czechach, są tylko możliwe przy pomocy oalej maszyn państwowe j, gdyby państwo uznało za wskazane użyć całej swej siły dla przyspieszenia czechizacji niemieckich części Czech.

Od 1897 przechodzi państwo straszne przesilenie. Austrya ginie, pokazało się niezbitcie, że Niemcy udusić się nie dadzą—trzeba było zawrócić z fałszywej drogi i zaczęło się zanosić na jakies ugodowo załatwienie. Niemcy, obok rozmaitych przesadnych pretensyj, które dało się obciąć, sami ofiarują język

ciunek i zważyć, czy żądane od nas ofiary stoją w jakimkolwiek stosunku do korzyści narodowych, politycznych, materialnych.

Przekroczenie tej granicy dla fałszywej polityki, dla jakichś Widoków słowiańskich, którym się trzeba strasznie dokładnie przypatrzeć, byłoby ze szkodą Polaków, prostem szaleństwem i zbrodnią, popełnioną na swoich własnych narodowych interesach nietylko w Austrii.

Nasze zaślepienie zdemoralizowało już Czechów w tym stopniu, że sobie z naszych zastrzeżeń nic nie robią. Jest to ostatni czas, żeby im powiedzieć: *sunt denique finis*.

Albo prawdą jest, że dla Polaków Austrya ma jeszcze jakąś wartość, albo to już stanowisko przedawnione. Jeżeli ma wartość jeszcze, a interes narodowy chciałby tę wartość zachować na dalszą przyszłość, to nie wolno dłużej igrać z ogniem, bo nam panowie Pacaki spalą dom.

Włości rentowe.

Rozpoznany tekst

lr, 191.

Lwdw. Środa, 25. kwietnia 1900.

Rok V.

PRENUMERATA WYNOŚI:

we Lwowie: miesięcznie 2 korony, za dwurazową dost. do doWu dopłaca ale 60 h.

na prowincyi:

rocznie 26 K. 401, z 2-krotną wys. 321. — i kwartał 6 „ 60 „ . 8 „ —
miesięcz. 2 „ 2 „ 70 „

W Niemczech: miesięcznie . 4 kor.
W innych krajach „ . 6 „

Przenumeratory *Słowa Polskiego* płacą sa *Wiek XX*, 1 kor.

Zażda. zmiana adresu 40 h.

sas*ś

Stalowa płyta.

Niezawodnie do interesujących postaci wśród czeskich polityków należy małutki, pękaty, czerwony, ruchliwy, o wieczne zachrypłym głósie, wiceprezes czeskiego klubu — poseł Pacak. Niegdyś, gdy nie konfiskowano gazet, ale wsadzano w Czechach redaktora do kozy, Pacak, który pisać nie umiał i nie pisał, był redaktorem od siedzenia w kozie i siedywał. Wyszedł potem na porządnego adwokata, robił czupumie politytę i został posłem. Sława jego jednak, która doprowadziła małutkiego Pacaka na krzesło wiceprezesa czeskiego klubu — to — mała broszurka, napisana w kąpielach morskich, w której zebrał materiał do rozwiązania sprawy językowej i naszkicował rozwiązanie tak — jak się stało czynem w rozporządzeniu hr. Badeniego.

Ten małutki, dobroduszny niby, filuterny, łatwo zapalny, ale zgodliwy Pacak, narobił już dużo kłopotów w Austrii. A dzisiejsze depesze z Pragi donoszą, że cisnął w czeskim Sejmie bombę, która ma rozarwać wszystkie dotychczas powiazane nitki uro-

Słowo Polskie

wychodzi 2 razy dziennie

broni się przed czechizacją. Nie wolno zapominać, że w tych centrach niemieckich — jest 75 powiatów sądowych, w których ludność czeska nie wynosi 3 procent, w 25 powiatach nie dochodzi jednego procentu.

W tych zamkniętych okręgach jest na 1,592 000 Niemców zaledwie 51.000-Czechów. Mimo niesłychanych rezultatów narodowego czeskiego odrodzenia, mimo imponującej żywotności, dzielności, ninożliwości, rozpiętej się po całym kraju i wylewającej rokrocznie ogromny nadmiar do prowincyj niemieckich, dalsze zdobycze na terytorium niemieckim w Czechach, są tylko możliwe przy pomocy oalej maszyn państwowe j, gdyby państwo uznało za wskazane użyć całej swej siły dla przyspieszenia czechizacji niemieckich części Czech.

Od 1897 przechodzi państwo straszne przesilenie. Austrija ginie, pokazało się niezbitcie, że Niemcy udusić się nie dadzą—trzeba było zawrócić z fałszywej drogi i zaczęło się zanosić na jakieś ugodowo załatwienie. Niemcy, obok rozmaitych przesadnych pretensyj, które dało się obciąć, sami ofiarują język

ciunek i zważyć, czy żądane od nas ofiary stoją w jakimkolwiek stosunku do korzyści narodowych, politycznych, materialnych.

Przekroczenie tej granicy dla fałszywej polityki, dla jakichś Widoków słowiańskich, którym się trzeba strasznie dokładnie przypatrzeć, byłoby ze szkodą Polaków, prostem szaleństwem i zbrodnią, popełnioną na swoich własnych narodowych interesach nietylko w Austrii.

Nasze zaślepienie zdemoralizowało już Czechów w tym stopniu, że sobie z naszych zastrzeżeń nic nie robią. Jest to ostatni czas, żeby im powiedzieć: *sunt denique finis*.

Albo prawdą jest, że dla Polaków Austrija ma jeszcze jakąś wartość, albo to już stanowisko przedawnione. Jeżeli ma wartość jeszcze, a interes narodowy chciałby tę wartość zachować na dalszą przyszłość, to nie wolno dłużej igrać z ogniem, bo nam panowie Pacaki spalą dom.

Włości rentowe.

Podział na warstwy

- W zdigitalizowanym obiekcie ważne jest zawarcie obrazu, a nie jedynie rozpoznanego tekstu
- Obraz tekstu może być czarnobiały (tzw. maska)
- Czarnobiały plik w rozdzielczości 300DPI zajmuje znacznie mniej miejsca
- Do czarnobiałego pliku można dodać uproszczone przedstawienie kolorowego tła (np. poprzez podzielenie DPI 3x, 6x, itd.)
- Można dołączyć również informacje o kolorze źródłowym pikseli czarnych, czyli uzupełnienie maski
- Uzupełnienie maski (*foreground*) może mieć jeszcze większy dzielnik DPI (12-16x)

MRC

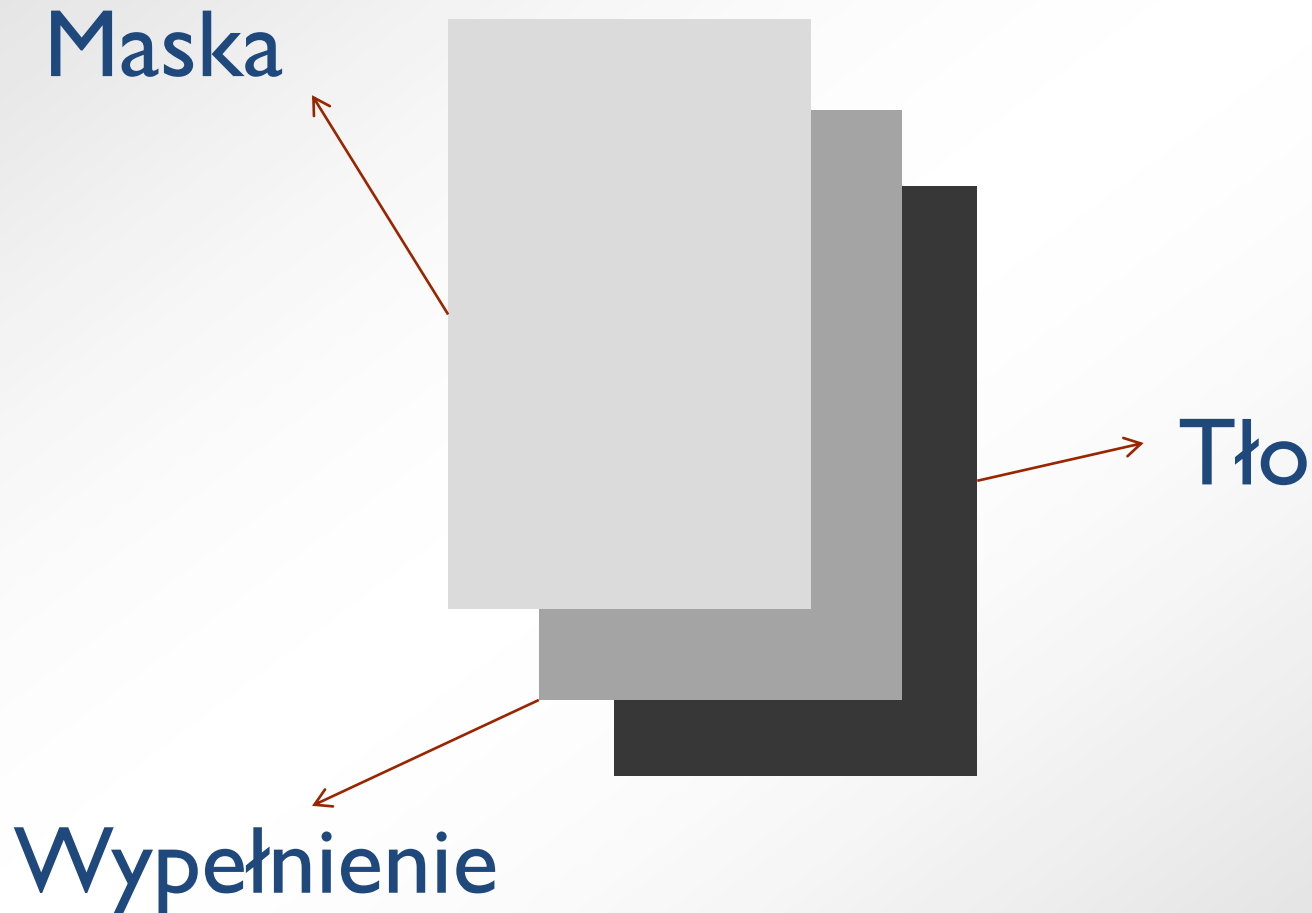
Obraz dokumentu może się więc składać z trzech podstawowych warstw:

- Maska (czarnobiałe przedstawienie treści)
- Uzupełnienie maski kolorem (*foreground*)
- Tło (*background*)

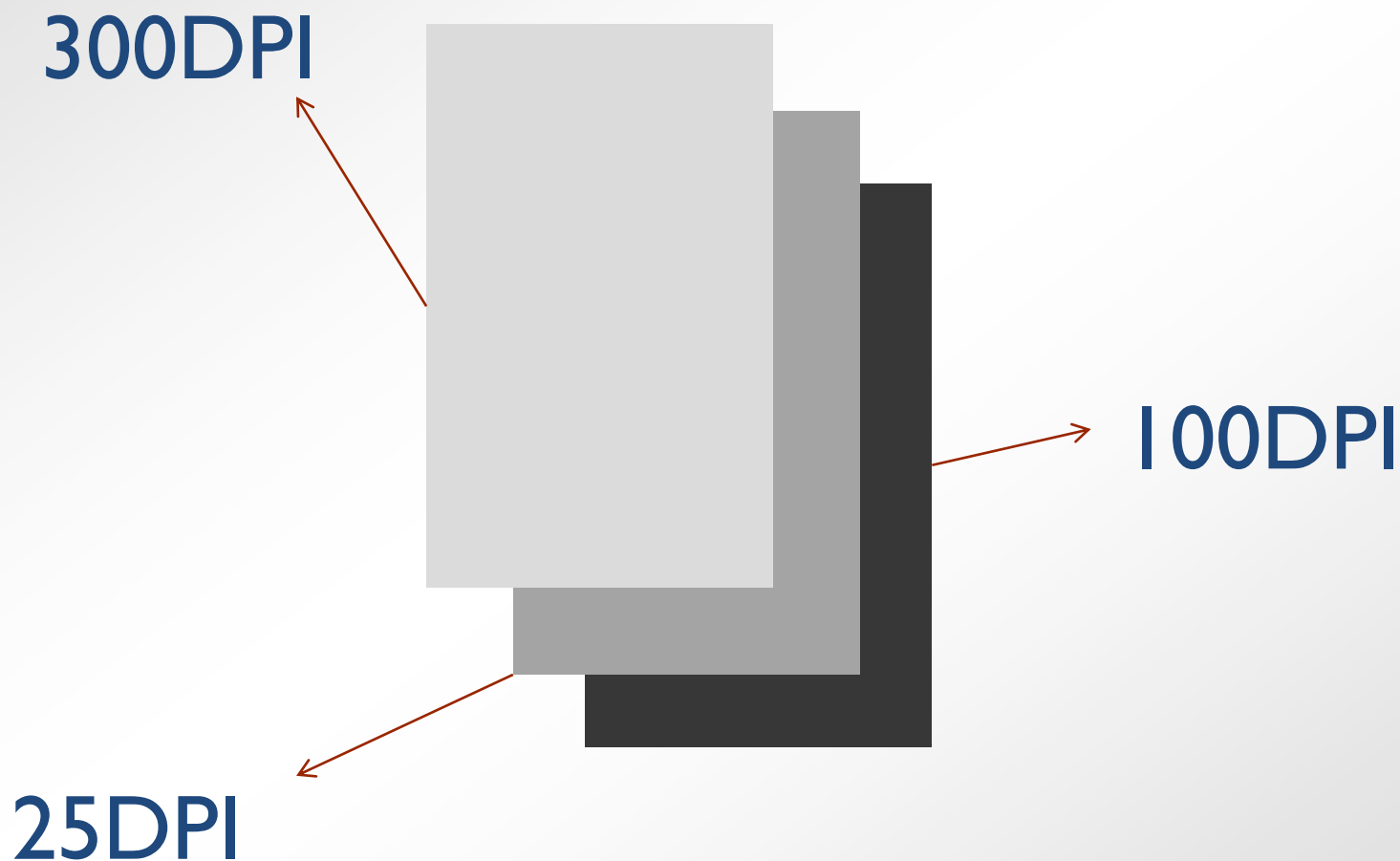
Każdy z obrazów dodatkowo może być kodowany stratnie za pomocą zaawansowanych algorytmów

Taki sposób kompresji nazywa się Mieszaną Zawartością Rastrową (MRC – *Mixed Raster Content*), lub kompresją adaptacyjną

Schemat MRC

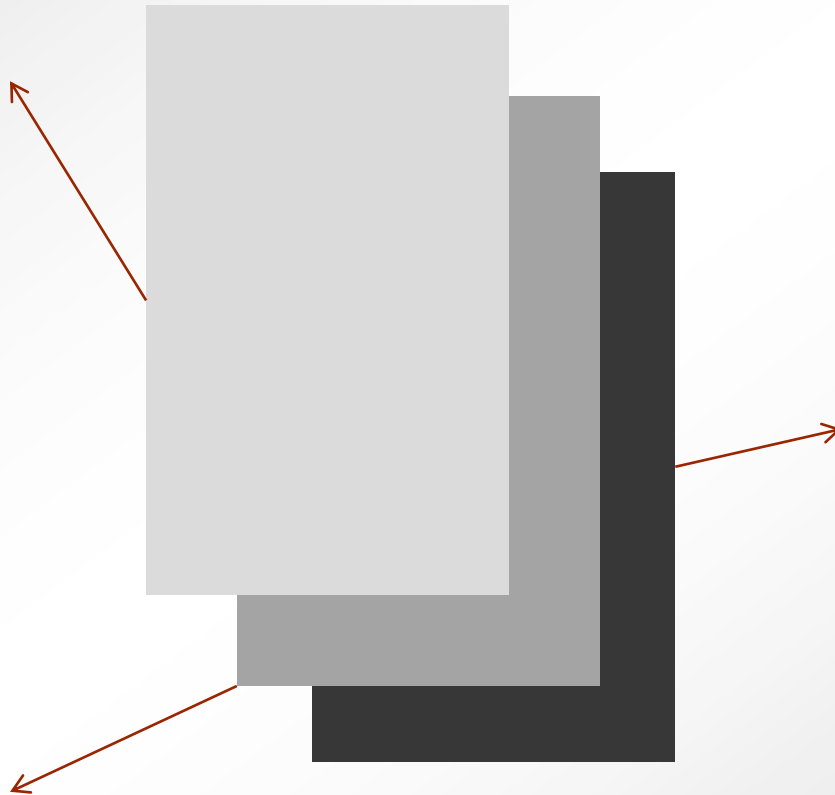


Dzielnik DPI



Kolor

1 BIT



16 BIT

4 BIT

Obliczenia (przy kolorze 24BIT)

- Maska = 300DPI RGB – 300DPI B&W: 4,16%
- Wypełnienie = 300DPI – 25DPI: 0,69%
- Tło = 300DPI – 100DPI: 11,11%

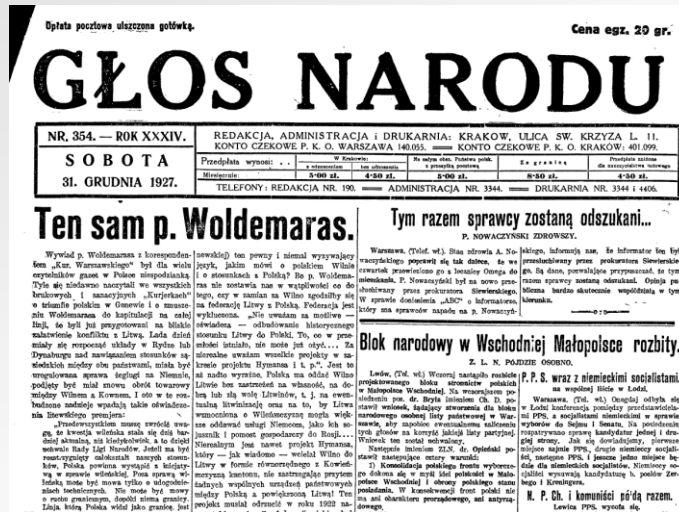
- Łącznie – $4,16\% + 0,69\% + 11,11\% = 15,96\%$

Obliczenia (zmienny kolor)

- Maska = 300DPI RGB – 300DPI B&W: 4,16%
- Wypełnienie = 300DPI – 25DPI 4BIT: 0,12%
- Tło = 300DPI – 100DPI: 5,56%

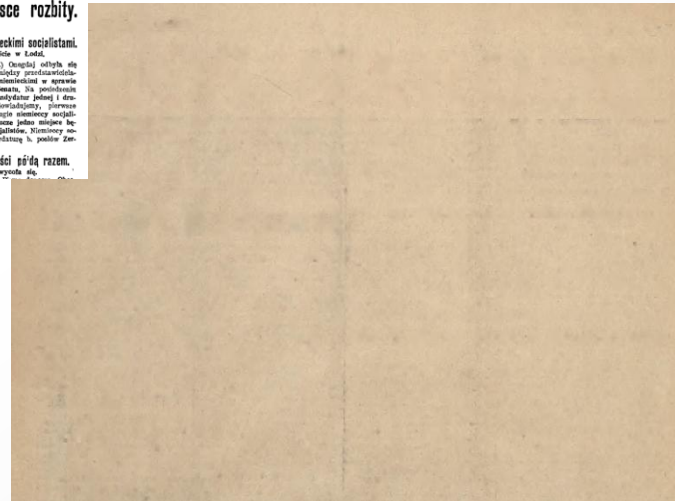
- Łącznie – $4,16\% + 0,12\% + 5,56\% = 9,84\%$

Inne zastosowanie MRC



Maska

Tłó



Inne zastosowanie MRC

Opłata pocztowa uliszczona gotówką. Cena egz. 20 gr.

GŁOS NARODU

NR. 354. — ROK XXXIV. SOBOTA 31. GRUDNIA 1927.

REDAKCYJA, ADMINISTRACJA I DrukARNIA: KRAKÓW, ULICA SW. KRZYŻA L. 11. KONTA CZEKOWE P. K. O. WARSZAWA 140.055. KONTA CZEKOWE P. K. O. KRAKÓW: 401.009.

Przebiegła wysyłka: ..	z przesyłką	z przesyłką pocztową	z przesyłką pocztową	z przesyłką pocztową
4.000 zł.	4.240 zł.	4.300 zł.	4.360 zł.	4.420 zł.

TELEFONY: REDAKCYJA NR. 190. ADMINISTRACJA NR. 3344. DrukARNIA NR. 3344 I 4406.

Ten sam p. Woldemaras.

Wypiód p. Woldemarasa z koresponden-
cją „Kur. Warszawskiego” był dla wielu
sympolików gwałt w Polsce siemopodlaska
Tyle się niekiedy znaczący w wywiadach do
brakowych i nasazyjnych „Kurjerkach”
o siamie polskich w Gwernii i o mianie
nie Woldemarasa do kapitanów na cze-
li, że był jui przegrywani na Miłkie
malowanie konfliktu z Litwą. Lada kied
miałby się rozpocząć układy w Rybnie hab
Dymalnego nad nawiazaniem stosunków z
skłódkich między nim potawianami, nada był
wyprowadzona sprawa legacji na Niemca,
podjęty był mial nowa obiet twarzowy
między Wilnem a Kownem. I oto w te roz-
budzone nadzieje wpełniał także odwiecz-
nie literackie opowiadanie.

„Jednomyślność naszą wywołał wpa-
g, że kwestie wielokrotnie się dotąd ha-
dowało, nie śladów, a to do dziś
wielkie kłopoty Litwy. Jakiś na był
zastrzeżony nakładami naszych stano-
wów. Polska powinna wyrazić si history-
czną, Litwa powinna wyrazić si history-
czną w sprawie wielokrotnie. Pora sprawa W-
liska może być nowa tyła o słabości
złach bezczelnych. Nie może być nowy
o rona przetrzymać, dopóki sama granicy.
Lada kied Polska widzi jaka granicy, jest

Tym razem sprawy zostaną odszukani..

Warszawa, (Telef. wł.) Słuchaliśmy A. No-
waczynskiego, który mówi o polskim Włcha
i o stosunkach z Polską. Do p. Woldemar-
asa nie należy nie w wywiadach do do-
lego, czy w samian na Wilno spodziły się
na federacji Litwy z Polską. Podkreśla jui
wystawiona. „Nie uważam na motiwe
— odwołanie historycznego
stosunku Litwy do Polski. To, co w pro-
mohi intubacji, nie może jui być... Za
siemolka swian wniknie projekt w na-
kresie projektu Hymanusa i L. P.”. Jest to
naślad wyrażenia. Polska ma oddać Wilno
Litwie bez narzutek na własność, na do-
lona hab się woli Litwinów, L. J. na ewen-
tualną litwinizację ona na to, by Litwa
wznowiona, o Włchowskiej może wiele
się oddawać usługi Niemcom, jako kł
janki i pomost gospodarczy do Angli...
Władymir jest nawet projekt Hymanusa
litwy — jak wiadomo — wolał Wilno do
Litwy w formie oświeceniowej z Kowal-
mizacją kontem, nie zastrzeżonej przysię-
gą między Polską a powstaniem Litwy. Ten
projekt miał odrośnięć w roku 1922 na-

Wypiód, informacja, że informacja ten był
przebiegła wysyłka...
do siamie polskich w Gwernii i o mianie
nie Woldemarasa do kapitanów na cze-
li, że był jui przegrywani na Miłkie
malowanie konfliktu z Litwą. Lada kied
miałby się rozpocząć układy w Rybnie hab
Dymalnego nad nawiazaniem stosunków z
skłódkich między nim potawianami, nada był
wyprowadzona sprawa legacji na Niemca,
podjęty był mial nowa obiet twarzowy
między Wilnem a Kownem. I oto w te roz-
budzone nadzieje wpełniał także odwiecz-
nie literackie opowiadanie.

Blok narodowy w Wschodniej Małopolsce rozbity.

Warszawa, (Telef. wł.) Wschodni nastąpiło powstanie
z Małopolski Wschodniej. Na wywołanie po-
wołania p. dr. Była imieniem Ch. D. po-
sady wolał, kładący straszenie dla kłosa
narodowego odwołany litwy potawian w Wa-
sawie, aby napoleo ewentualnie talowano
tych gwałt na korespondentów litwy potawian.
Wskazano ten został sekwencyjny.
Kwestie imienia TSK, opowiada jui
stał zastępczo strony warunki.
2) Komitety polskie brama wybra-
ny do kłosa, aby w spłi kłsi polskimi z Ma-
łopolski Wschodniej i obywateli polskie sta-
powstał. W konwencji front polski się
na się skłonił porządkiem, nie zastrze-
żony.

P. P. S. wraz z niemieckimi socjalistami.
na wojnie litwy z Łódź.
Warszawa, (Telef. wł.) Oczekiwać się
w Łodzi konferencji politycy przedstawicieli
ni PPS, a socjalistami orientacji o sprawie
wyborów do Sejmu i Senatu. Na posiedze-
niach rozstrzygnięto sprawę kandydatury jui i do-
głębli strony. Jak się dowiadujemy, pierwsze
podjęto najnie PPS, drugie niemieckie socja-
listy, następnie PPS. I jeszcze jedno ważne be-
dzia dla Niemców socjalistami. Kwestory na-
głębli wyrażali kandydaturę jui powoła-
nie Kowalczaka.

M. P. Ch. i komuniści sądzą razem.
Lewica PPS, wysłała się.

Opłata pocztowa uliszczona gotówką. Cena egz. 20 gr.

GŁOS NARODU

NR. 354. — ROK XXXIV. SOBOTA 31. GRUDNIA 1927.

REDAKCYJA, ADMINISTRACJA I DrukARNIA: KRAKÓW, ULICA SW. KRZYŻA L. 11. KONTA CZEKOWE P. K. O. WARSZAWA 140.055. KONTA CZEKOWE P. K. O. KRAKÓW: 401.009.

Przebiegła wysyłka: ..	z przesyłką	z przesyłką pocztową	z przesyłką pocztową	z przesyłką pocztową
4.000 zł.	4.240 zł.	4.300 zł.	4.360 zł.	4.420 zł.

TELEFONY: REDAKCYJA NR. 190. ADMINISTRACJA NR. 3344. DrukARNIA NR. 3344 I 4406.

Ten sam p. Woldemaras.

Tym razem sprawy zostaną odszukani..

Warszawa, (Telef. wł.) Słuchaliśmy A. No-
waczynskiego, który mówi o polskim Włcha
i o stosunkach z Polską. Do p. Woldemar-
asa nie należy nie w wywiadach do do-
lego, czy w samian na Wilno spodziły się
na federacji Litwy z Polską. Podkreśla jui
wystawiona. „Nie uważam na motiwe
— odwołanie historycznego
stosunku Litwy do Polski. To, co w pro-
mohi intubacji, nie może jui być... Za
siemolka swian wniknie projekt w na-
kresie projektu Hymanusa i L. P.”. Jest to
naślad wyrażenia. Polska ma oddać Wilno
Litwie bez narzutek na własność, na do-
lona hab się woli Litwinów, L. J. na ewen-
tualną litwinizację ona na to, by Litwa
wznowiona, o Włchowskiej może wiele
się oddawać usługi Niemcom, jako kł
janki i pomost gospodarczy do Angli...
Władymir jest nawet projekt Hymanusa
litwy — jak wiadomo — wolał Wilno do
Litwy w formie oświeceniowej z Kowal-
mizacją kontem, nie zastrzeżonej przysię-
gą między Polską a powstaniem Litwy. Ten
projekt miał odrośnięć w roku 1922 na-

Blok narodowy w Wschodniej Małopolsce rozbity.

Warszawa, (Telef. wł.) Wschodni nastąpiło powstanie
z Małopolski Wschodniej. Na wywołanie po-
wołania p. dr. Była imieniem Ch. D. po-
sady wolał, kładący straszenie dla kłosa
narodowego odwołany litwy potawian w Wa-
sawie, aby napoleo ewentualnie talowano
tych gwałt na korespondentów litwy potawian.
Wskazano ten został sekwencyjny.
Kwestie imienia TSK, opowiada jui
stał zastępczo strony warunki.
2) Komitety polskie brama wybra-
ny do kłosa, aby w spłi kłsi polskimi z Ma-
łopolski Wschodniej i obywateli polskie sta-
powstał. W konwencji front polski się
na się skłonił porządkiem, nie zastrze-
żony.

M. P. Ch. i komuniści sądzą razem.
Lewica PPS, wysłała się.

Wyświetlenie

Nie można zastosować modelu MRC bez maski

Tłó



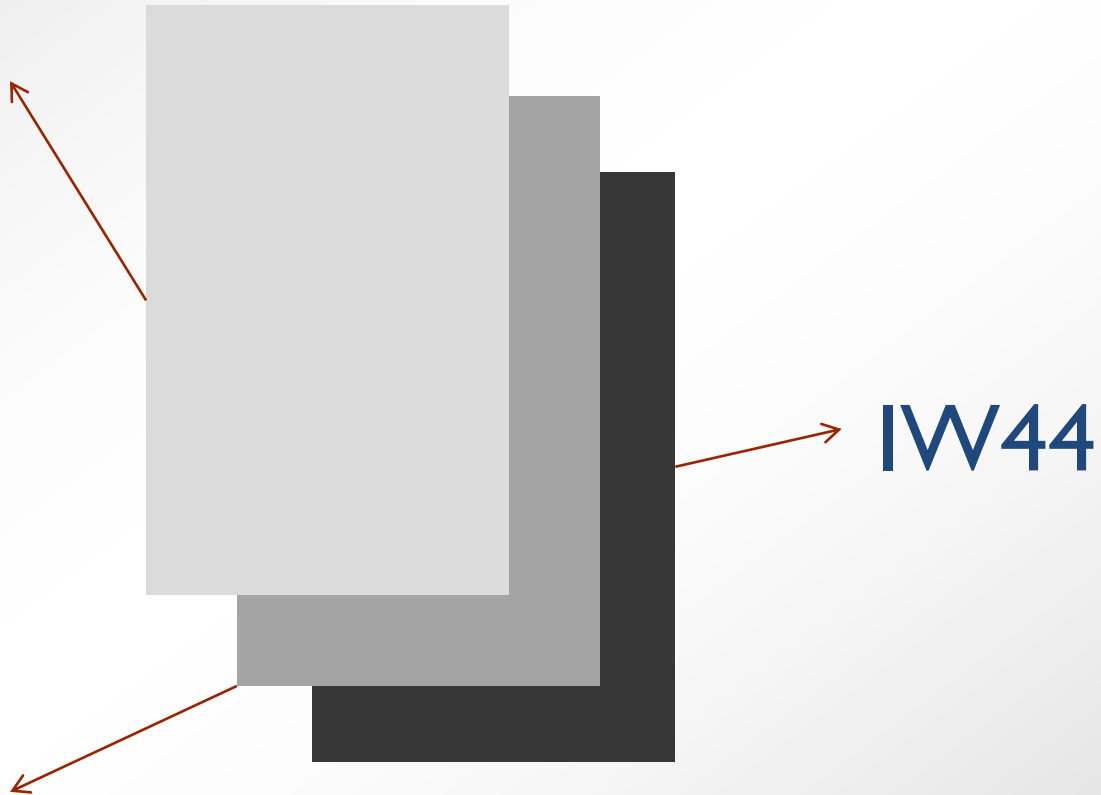
Wypełnienie

MRC

- Format, który w modelowy sposób wykorzystuje kompresję MRC to DjVu
- DjVu pozwala również na dodanie do dokumentu rozpoznanego tekstu (poprzez proces OCR)
- W formacie zaaplikowano również kompresję pomiędzy stronami (słownik kształtów)
- Na potrzeby formatu zastosowano dwa modele kompresji obrazu, JB2 i IW44

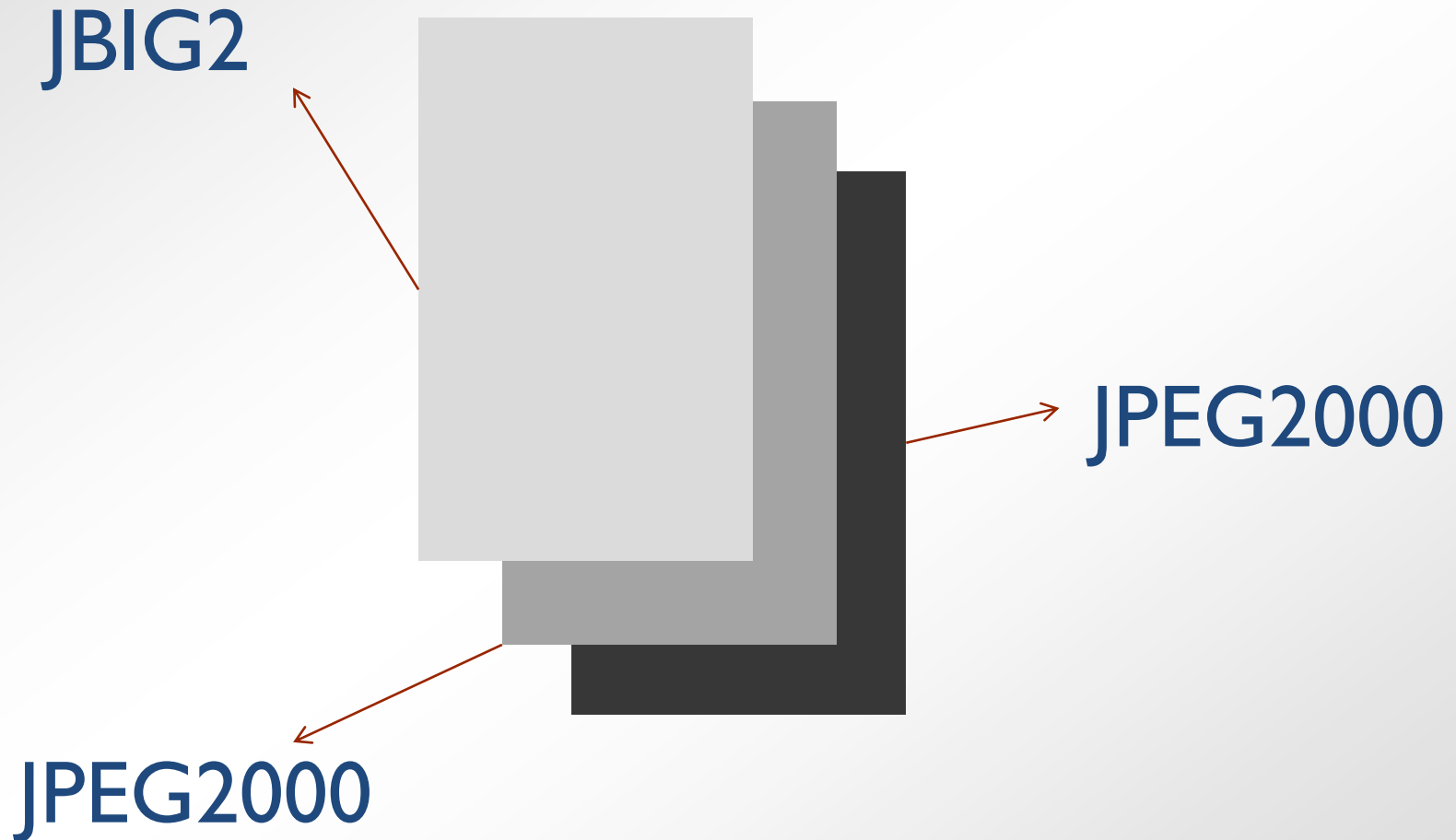
Komrpesja w DjVu

JB2



IW44

Komrpesja w PDF 1



Komrpesja w PDF 2

JBIG2



JPEG2000

Sposoby kompresowania

- JB2/JBIG2 kompresuje stratnie, wykorzystując kompresowanie algorytmiczne i kodowanie blokowe
- IW44/JPEG2000 wykorzystuje wielopoziomowe falki
- IW44 jest algorytmem stworzonym specjalnie na potrzeby formatu DjVu, JPEG2000 jest formatem niezależnym
- Przy zastosowaniu kompresji falkami w formacie DjVu istnieje możliwość wykorzystania słownika kształtów pozwalającego dodać jeden słownik do całego dokumentu
- Jeden słownik dla dokumentu wpływa nieznacznie na kompresję, jedynie poprawia szybkość dekodowania
- Słownik wykorzystywany dla kilku plików uniemożliwia otworzenie obrazu bez słownika

Efekty kompresji

- Model kompresji MRC jest zorientowany na segmentację obrazu
- W przypadku kompresji bez segmentacji cały proces odbywa się w pełni automatycznie (program dobiera samodzielnie wyrażenia słownikowe, wielkość bloku ustalona jest dla całego obrazu, falki wykorzystują jedną funkcję macierzystą)
- Segmentacja wymaga ustawienia filtrów, które wybiorą z obrazu treść i stworzą na tej podstawie maskę
- Tworzenie maski wymaga więc kontroli i dostosowywania segmentacji do konkretnego przypadku

Filtry w DjVu

- Treshold
- Shape-filter
- Pix-filter
- Inversion-level
- Block-size
- Background-Subsample
- Foreground-Subsample

Nie używane:

- Ihibit-Foreback-Level; high-variation-foreground; thickening; render-size; blurring-size, block-overlap; upsample; subsample; edge-size; pages-per-dict; subsample-refine

Konwersja do formatu DjVu w programie ABBYY FineReader 11

- Oprogramowanie FineReader pozwala wykonać automatycznie cały proces segmentacji, użytkownik nie ma wpływu na proces
- Czy jest to zgodne z założeniami modelu MRC?
- Brak wpływu na segmentacje w programie FineReader wpływa negatywnie na tworzenie obrazu
- Brak zrozumienia modelu kompresji MRC często daje efekt źle wytworzonych plików DjVu

DjVu wykonane programem ABBYY FineReader 11

Należność pocztowa uliszczona gotówką. Cena egz. 25 gr.

GŁOS NARODU

NR. 1. — ROK XXXVIII	REDAKCJA, ADMINISTRACJA I DrukARNIA KRAKÓW, ULICA SW. KRZYŻA L. II. KONTA CZEKOWE P. K. O. WARSZAWA 140.055 — KONTA CZEKOWE P. K. O. KRAKÓW 401.099.					
CZWARTEK	Przedpłata wynosi:	z KRAKOWA	z WARSZAWY	z PRAGI	Przedpłata przelana dla abonamentów zagranicą	Za każdy emblema adreśu dopłata 50 gr.
1 STYCZNIA 1931.	6*20 zł.	6*20 zł.	6*20 zł.	6*20 zł.	6*20 zł.	
TELEFONY: REDAKCJA Nr. 101-90, ADMINISTRACJA Nr. 133-44, DRUKARNIA Nr. 133-44 i 144-06. ODDZIAŁ LWJ W BRZOZDKA 2 B.						

WINA na okres świąteczny
10% taniej

polecac:
WOJCIECH CŁSOWSKI, Kraków, Mały Rynek.

Na zabawy i rauty daje wina, wódki i likiery w komis.

Porachunek!

Zgany dziś róz, będą najnamuliejzy i najolejczy od rzazu obfaluwy panstwa. Nie znajdujemy w nim ani jednego wydażca, rzazu wielkiej wagi, którego wspomnienie mogłoby dziś ogrzeć polskie serce. Nie dostrzegamy w takiej dziedzinie pracy narodowej poważniejszący wyników dodatnich. Czy rzadimy okłóm na życie gospodarza, które analizują się w pelnym kryzysie i produkcji czasu, więcej bezroboczych, czy...

liczeństwo polskie, zsiemralizowane nie- wola, jest — jeśli chodzi o jego warunki inteligentne i kierownicze — o wiele mniej skuteczny i rozumny, niż przedstawiała je sobie nasza młodość wiana. Wylęcany dzie- niąc zachodnia, moralnie zdrowa i tęg- a sportrzestany, że poza się pokolónie współ- czesne wroci w życie publiczne dano więcej karyerowiczostwa, niż przekonał, dano wię- cej głęboży połączonej z tupetem, niż wie- dzi i powagi. Dzierżypją się warstwa in-

Polecamy!

po nalkających ce- nach i szybkich w wielkim wyborze

Linoleum, Ceraty, Dywany welniane i pluszowe Chodniki, Kapy na łózka, Koce i Pledy, Narzuty, Firanki, Portjery, Chodniki kokosowe Płaszczki gumowe i impregnowane



PRZEMYSŁ-LINOLEUM
Kraków. Rynek 10.

Porachunek!

Zgany dziś róz, będą najnamuliejzy i najolejczy od rzazu obfaluwy panstwa. Nie znajdujemy w nim ani jednego wydażca, rzazu wielkiej wagi, którego wspomnienie mogłoby dziś ogrzeć polskie serce. Nie dostrzegamy w takiej dziedzinie pracy narodowej poważniejszący wyników dodatnich. Czy rzadimy okłóm na życie gospodarza, które analizują się w pelnym kryzysie i produkcji czasu, więcej bezroboczych, czy...

liczeństwo polskie, zsiemralizowane nie- wola, jest — jeśli chodzi o jego warunki inteligentne i kierownicze — o wiele mniej skuteczny i rozumny, niż przedstawiała je sobie nasza młodość wiana. Wylęcany dzie- niąc zachodnia, moralnie zdrowa i tęg- a sportrzestany, że poza się pokolónie współ- czesne wroci w życie publiczne dano więcej karyerowiczostwa, niż przekonał, dano wię- cej głęboży połączonej z tupetem, niż wie- dzi i powagi. Dzierżypją się warstwa in-

Cena egz. 25 gr.

GŁOS NARODU

DRUKARNIA KRAKOW, ULICA SW. KRZYŻA L. II. 140.055 — KONTA CZEKOWE P. K. O. KRAKÓW 401.099.			
Przedpłata przelana dla abonamentów zagranicą	Za każdy emblema adreśu dopłata 50 gr.		
6*20 zł.	6*20 zł.	6*20 zł.	6*20 zł.
.133-44, DRUKARNIA Nr. 133-44 i 144-06. ODDZIAŁ LWJ W BRZOZDKA 2 B.			

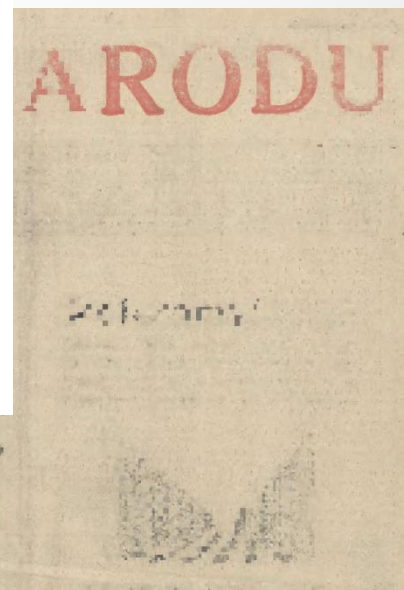
Polecamy!

po nalkających ce- nach i szybkich w wielkim wyborze

Linoleum, Ceraty, Dywany welniane i pluszowe Chodniki, Kapy na łózka, Koce i Pledy, Narzuty, Firanki, Portjery, Chodniki kokosowe Płaszczki gumowe i impregnowane



PRZEMYSŁ-LINOLEUM
Kraków. Rynek 10.



Zalety dobrej segmentacji

- Przy dobrze wykonanej segmentacji można wykorzystać zalety modelu kompresji MRC
- Główną zaletą jest dzielenie DPI obrazu
- Jeżeli segmentacja jest wykonana niepoprawnie, by nie tracić jakości obrazu dzielnik DPI nie może być duży
- Niepoprawna segmentacja treści sprawia, że tekst jest nieczytelny
- W dalszej konsekwencji optyczne rozpoznanie tekstu jest nieskuteczne

Niepoprawna segmentacja

nieśli, że właściele tego gr
ie
szczer
żep k ni
że
U
Bie
ł
ane
g a wię
po
o tego
w pierwszym un
Sejm stępuj
ości gminne
re tacy
et.
Ulanowie, Bielinach i udr

nieślono, że właściele tego gr
ie ze swego mienia wywłaszczeni
ni, dzisiaj stali się żebrakami
em jest w tej petycyi, że drog
Ulanowa do Bielin jest w te
ędzy drogą a rzeką Sanem
oko 2 metry. Jest nagła wię
ym właścicielom przyjąć w pom
obie uczynić poprawkę do tego
posób, ażeby w pierwszym pun
isyi po słowach: „Sejm odstępuj
ości gminnej Ulanowa,“ było
rezentacyi powiatu Niska do
et. w przedmiocie zabezpieczenia
Ulanowie, Bielinach i Rudni

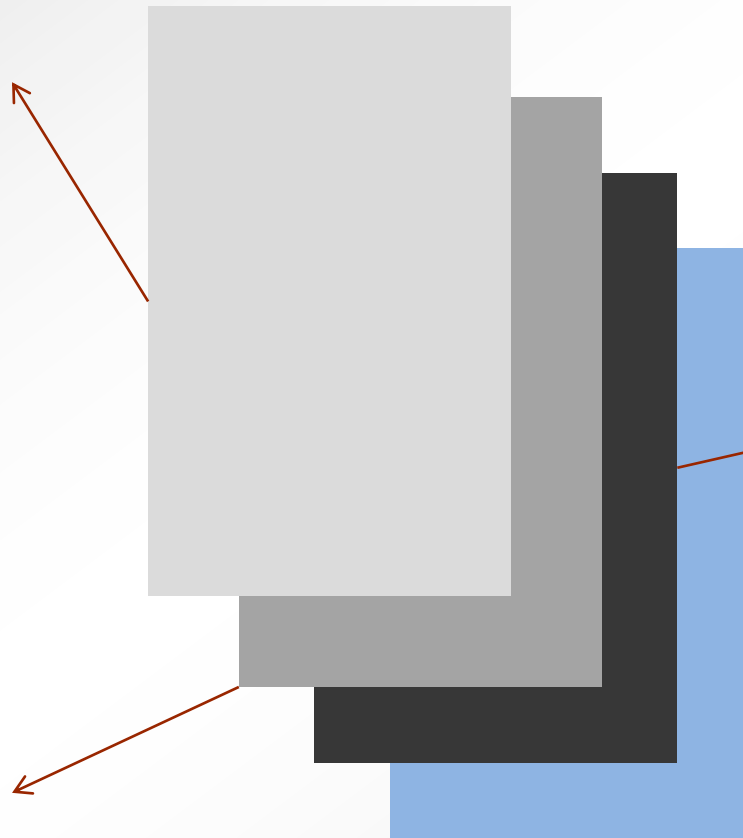
Zarządzanie materiałem w formacie DjVu

- DjVu jest formatem przeznaczonym do kompresji dokumentów
- Pojedynczy obraz tworzy stronę dokumentu i jest zapisywany osobno
- Dla wszystkich plików tworzony jest dodatkowy plik indeksu
- Indeks pozwala zarządzać stronami w dokumencie
- DjVu daje możliwość jakby włączenia wszystkich stron do indeksu, wówczas dokument zawarty jest w jednym pliku
- Wiele plików nazywamy zapisem rozdzielonym (*indirect*), jeden plik nazywamy zapisem scalonym (*bundle*)

- W formacie DjVu istnieje możliwość zawarcia warstwy rozpoznanego tekstu, który jest umieszczony pod widocznym obrazem strony

Komprpesja w DjVu

JB2



IW44

Tekst
(OCR)

IW44

OCR

- OCR – *Optical Character Recognition*
- Optyczne rozpoznanie tekstu

Obecnie wykorzystuje się trzy główne silniki OCR:

- Silnik firmy ABBYY (FineReader)
 - Tesseract (darmowy, Google)
 - Read I.R.I.S. (Document Express Enterprise)
-
- Proces OCR Nigdy nie jest wykonany w 100% poprawnie

Priorytety procesu OCR

- Jedynie silnik firmy ABBYY pozwala na odpowiednie wpływanie na rozpoznaną strukturę dokumentu
- Rozpoznanie struktury dokumentu jest dla przyszłości bibliotek cyfrowych najważniejszym czynnikiem oceny jakości OCR*
- Silniki ABBYY i Tesseract pozwalają na douczanie programu rozpoznającego
- Wykonanie procesu OCR można przeprowadzić na plikach czarnobiałych (Read I.R.I.S, Tesseract) albo kolorowych (ABBYY)
- Silnik ABBYY również konwertuje obraz rozpoznawany do czerni i bieli

* Mühlberger, Günter. „Digitalisierung historischer Zeitungen aus dem Blickwinkel der automatisierten Text- und Strukturerkennung (OCR)”. Zeitschrift für Bibliothekswesen und Bibliographie 58, nr 1 (2011): 10–18.

