

ITALIAN PATENT OFFICE

Document No.

102010901828322A1

Publication Date

20111012

Applicant

SISVEL TECHNOLOGY S.R.L.

Title

METODO PER LA GENERAZIONE E RICOSTRUZIONE DI UN FLUSSO
VIDEO STEREOSCOPICO COMPATIBILE E RELATIVI DISPOSITIVI DI
CODIFICA E DECODIFICA.

Descrizione dell'Invenzione Industriale dal titolo:

-SVT007-

“METODO PER LA GENERAZIONE E RICOSTRUZIONE DI UN FLUSSO VIDEO STEREOSCOPICO COMPATIBILE E RELATIVI DISPOSITIVI DI CODIFICA E DECODIFICA”

di Sisvel Technology, di nazionalità italiana, con sede in Via Castagnole 59, 10060 None (TO), ed elettivamente domiciliata presso i Mandatari Ing. Roberto Dini (No. Iscr. Albo 270 BM), Ing. Marco Camolese (No. Iscr. Albo 882 BM), Ing. Antonio Di Bernardo (No. Iscr. Albo 1163 BM), Dott. Giancarlo Reposio (No. Iscr. Albo 1168 BM) ed Ing. Andrea Grimaldo (No. Iscr. Albo 1060 BM) c/o Metroconsult S.r.l., Via Sestriere 100, 10060 None (TO).

Inventori designati:

- BALLOCCA Giovanni, Via Lavazza 56/8, 10131 Torino

- D'AMATO Paolo, Via Costabella, 21, 00195 ROMA

Depositata il

No.

DESCRIZIONE

La presente invenzione si riferisce ad un metodo di generazione di un flusso video stereoscopico digitale, ossia di un flusso video che, elaborato in modo opportuno in un dispositivo di visualizzazione, produce sequenze di immagini del tipo che vengono percepite come tridimensionali da un osservatore.

Tale metodo consente inoltre all'utente che dispone di un decoder e di un televisore convenzionale (non stereoscopico) la visualizzazione delle suddette immagini in 2D.

La presente invenzione si riferisce inoltre ad un dispositivo di codifica e ad un dispositivo di decodifica di un flusso video stereoscopico in forma digitale.

Nel corso degli ultimi anni, il mondo della produzione cinematografica, stimolato

dalla disponibilità di nuovi strumenti di produzione resi possibili dall'avvento delle tecnologie digitali, ha rivolto grande attenzione e ingenti risorse alla produzione di contenuti 3D di tipo stereoscopico.

L'interesse per il 3D si sta estendendo al possibile uso domestico, cioè con visualizzazione delle immagini su di un televisore. Ad esempio, alcuni operatori di pay-TV stanno per iniziare a trasmettere programmi in 3D.

L'approccio più comune nella presentazione di contenuti video stereoscopici è quello basato sulla visualizzazione di due flussi video indipendenti, destinati rispettivamente all'occhio destro e all'occhio sinistro, che vengono ricomposti dal cervello nella rappresentazione di un oggetto tridimensionale.

I contenuti stereoscopici per uso domestico sono generalmente contenuti video in alta definizione e possono essere distribuiti su un supporto di memorizzazione di massa (DVD, Blu-ray disc, supporti magneto-ottici o a stato solido e così via) oppure attraverso canali di broadcasting (cablati o wireless) o via rete telematica (IP).

In ambiente di produzione, però, può risultare impossibile, con le strutture esistenti, trasferire e gestire in modo separato i due flussi prodotti dalle telecamere stereoscopiche che riprendono la scena da due punti di vista separati.

Inoltre, le reti di distribuzione che permettono di raggiungere l'utente finale hanno un dimensionamento tale da rendere antieconomico l'uso di due flussi ad alta definizione indipendenti per l'erogazione di un singolo servizio. Di conseguenza, è necessario adottare alcuni accorgimenti in fase di produzione che consentano di ridurre il bit-rate necessario al trasferimento ed alla fruizione dei contenuti.

Poiché i risultati degli studi sulle differenze nella percezione dei dettagli nelle immagini bidimensionali e tridimensionali sembrano indicare che, anche

abbassando la risoluzione dei contenuti stereoscopici rispetto a quella dei contenuti bidimensionali, la qualità percepita dall'utente rimane accettabile, sono state sviluppate tecniche differenti per l'impacchettamento delle due immagini componenti la vista stereoscopica in un singolo fotogramma.

Ad esempio, nel caso di un singolo fotogramma C ad alta definizione (1920 x 1080 pixel), vengono acquisite le due immagini che compongono il canale sinistro e quello destro (nel seguito L e R) con una risoluzione orizzontale pari alla metà della risoluzione di un fotogramma ad alta definizione e in seguito disposte fianco a fianco in un singolo fotogramma (formato side-by-side), come rappresentato in Figura 1a.

In questo modo, è possibile utilizzare un singolo flusso ad alta definizione per il trasporto dei due canali video indipendenti; al momento della decodifica i due semifotogrammi vengono separati e riportati al formato 16/9 applicando opportune tecniche di interpolazione.

Allo stesso modo, è possibile applicare un procedimento analogo dimezzando la risoluzione verticale, mantenendo intatta quella orizzontale, e disponendo i due fotogrammi L ed R uno sopra l'altro (formato top/bottom), come rappresentato in Figura 1b.

Il flusso video stereoscopico costituito dai fotogrammi composti viene quindi compresso, al fine di ridurre il bit-rate di trasporto per la distribuzione su rete di broadcasting, rete IP o supporto di memorizzazione di massa.

La tecnica di compressione attualmente più utilizzata nella distribuzione di video ad alta definizione è quella definita nello standard H.264/AVC.

I televisori ad alta definizione attualmente disponibili in commercio sono dotati di decoder H.264/AVC che supportano la decodifica dei formati fino al formato

1080p.

Uno dei requisiti sui quali si concentra l'attenzione dei fornitori di servizio, in particolare i broadcaster responsabili del servizio pubblico, è quello della retrocompatibilità dei segnali stereoscopici.

Infatti, per consentire agli utenti che già possiedono un decoder ad alta definizione di fruire dei servizi trasmessi, è desiderabile che i programmi 3D possano essere visualizzati anche come programmi 2D. Allo stesso modo è desiderabile che un contenuto 3D su DVD, Blu-ray disc 3D o sito Internet possa essere visualizzato dai televisori, o monitor, sia 2D che 3D.

Questo risultato può essere raggiunto in due modi: trasmettendo simultaneamente la versione 2D e quella 3D di uno stesso programma, oppure adottando una opportuna tecnica di codifica del flusso stereoscopico.

La prima opzione comporta ovviamente uno spreco di banda, che gli operatori del servizio desiderano evitare per quanto possibile.

Per quanto riguarda la seconda opzione, sono note nell'arte tecniche per la generazione di flussi stereoscopici retrocompatibili.

Una di queste tecniche è legata all'applicazione delle cosiddette "mappe di profondità", descritte ad esempio nelle domande di brevetto statunitensi no. US 2002/0048395 e no. US 2004/0101043.

In pratica, al video a colori bidimensionale viene associato un segnale, in forma di video supplementare in bianco e nero, che trasporta le mappe di profondità. Un opportuno decodificatore può ricostruire un video stereoscopico a partire dai dati ricevuti. Tuttavia, questa tecnica soffre degli stessi problemi della succitata trasmissione in 2D ed in 3D dello stesso programma: è infatti necessario trasferire due segnali video in parallelo, con elevato dispendio in termini di bit-rate di

trasporto.

Un'altra tecnica di codifica del flusso stereoscopico di tipo retrocompatibile è ad esempio quella denominata "multiview".

Poiché le coppie di immagini destra e sinistra che compongono il flusso video stereoscopico sono caratterizzate da un elevato grado di somiglianza, si utilizzano le stesse tecniche di soppressione della ridondanza spazio-temporale impiegate nella codifica dei flussi bidimensionali. Infatti, sottratto un certo offset dovuto alla distanza geometrica dei punti di ripresa (la distanza interoculare), le differenze tra l'immagine destra e quella sinistra sono piccole.

Lo standard MPEG2 è stato esteso con una specifica supplementare chiamata Multi View Profile (MVP); allo stesso modo, il successivo standard H.264/AVC è stato esteso con l'inclusione della specifica Multi View Coding (MVC).

Carattere comune a queste due specifiche è l'uso della codifica video scalabile: il flusso video stereoscopico viene compresso in un layer di base (il flusso base 2D) più un enhancement layer, che trasporta la seconda vista. La sintassi del flusso codificato garantisce la decodificabilità del video 2D anche da parte di decoder di vecchia generazione, purché conformi agli standard MPEG2 o H.264/AVC.

Tuttavia, il bit-rate necessario alla codifica dei flussi stereoscopici in uno dei formati sopra descritti risulta, ad oggi, ancora troppo elevato per permetterne l'impiego in ambiente di broadcasting.

Scopo della presente invenzione è pertanto quello di indicare un metodo di generazione e ricostruzione di un flusso video stereoscopico digitale di tipo retrocompatibile che possa essere distribuito da un broadcaster impiegando la stessa banda necessaria per un flusso 2D.

Un ulteriore scopo della presente invenzione è quello di indicare un metodo di

generazione e ricostruzione di un flusso video stereoscopico digitale di tipo retrocompatibile basato sull'uso di fotogrammi compositi, che sia applicabile qualunque sia il modo in cui vengono impacchettate le immagini destra e sinistra all'interno di detti fotogrammi compositi.

Questi ed altri scopi della presente invenzione sono ottenuti mediante un metodo di generazione e ricostruzione di un flusso video stereoscopico incorporante le caratteristiche delle rivendicazioni allegate, le quali formano parte integrante della presente descrizione.

La presente invenzione descrive inoltre un dispositivo di codifica ed un dispositivo di decodifica di un flusso video stereoscopico, ed un flusso video stereoscopico.

L'idea generale alla base della presente invenzione è quella di indicare un metodo per generare e ricostruire un flusso video comprendente una pluralità di fotogrammi impacchettati secondo una tecnica di impacchettamento, in modo tale che il risultato della decodifica del flusso video sia fruibile sia da un decoder 2D che da un decoder 3D.

In una forma preferita di realizzazione, il metodo secondo l'invenzione consente di predisporre un flusso video stereoscopico in modo che tale flusso sia decodificabile da un decodificatore 2D conforme alla norma H.264/AVC (senza apportare alcuna modifica a detto decodificatore) e che sia possibile visualizzarlo su un display ad alta definizione.

Un apposito decodificatore, accoppiato ad un display stereoscopico, invece, permette la visione del flusso video stereoscopico in 3D.

La retrocompatibilità del flusso video stereoscopico è resa possibile grazie al diverso utilizzo, in fase di codifica, di metadati già presenti all'interno del flusso

video stereoscopico.

Tali metadati definiscono una regione del fotogramma composito contenente un'immagine, destra o sinistra, del flusso video stereoscopico.

Per un decodificatore 2D detti metadati contengono un'informazione che indica al decodificatore che, dopo avere decodificato il fotogramma, deve presentare in uscita solo una regione dell'immagine composita decodificata, ossia quella che contiene una sola immagine, destra o sinistra, del flusso video stereoscopico.

In questo modo, il broadcaster deve trasmettere un solo flusso video stereoscopico, il quale viene elaborato in modo opportuno dal decodificatore, sia esso abilitato o meno alla decodifica di flussi video stereoscopici: lo stesso flusso video può essere riprodotto in formato 2D o 3D a seconda delle caratteristiche del decoder che lo elabora.

Ulteriori scopi e vantaggi della presente invenzione appariranno maggiormente chiari dalla descrizione che segue di alcuni esempi di realizzazione forniti a titolo esplicativo e non limitativo.

Tali esempi di realizzazione vengono descritti con riferimento ai disegni allegati, in cui:

- le Figure 1a, 1b ed 1c rappresentano un fotogramma composito di un flusso video stereoscopico, rispettivamente in formato side-by-side, in formato top-bottom ed in un formato alternativo;
- la Figura 2 rappresenta uno schema a blocchi di un dispositivo per generare un flusso video stereoscopico secondo l'invenzione;
- le Figure 3a, 3b e 3c rappresentano i fotogrammi composti delle Figure 1a, 1b ed 1c, sui quali è evidenziata in tratteggio una finestra di delimitazione dell'immagine, o cropping rectangle;

- le Figure 4a e 4b rappresentano gli schemi a blocchi di un dispositivo per la fruizione di immagini televisive, ricevute sotto forma di un flusso video digitale, rispettivamente nel caso in cui il dispositivo consenta la visualizzazione di sole immagini 2D ed in quello in cui il dispositivo consenta anche la visualizzazione di immagini 3D.

Con riferimento alla Figura 2 viene illustrato uno schema a blocchi di un dispositivo 100 per generare un flusso video stereoscopico 101 secondo l'invenzione.

Il dispositivo 100 riceve in ingresso due sequenze d'immagini 102 e 103, ad esempio due flussi video, destinate rispettivamente all'occhio sinistro L ed all'occhio destro R.

Il dispositivo 100 permette di multiplexare due immagini delle due sequenze d'immagini 102 e 103. Esso comprende a tal fine un modulo di assemblaggio 105 in grado di inserire i pixel d'immagini ricevute in ingresso in un'unica immagine composita C.

Nel prosieguo della presente descrizione, quando si fa riferimento all'inserimento di un'immagine all'interno di un fotogramma composito o fotogramma contenitore C, s'intende fare riferimento ad una procedura per cui si genera, con mezzi hardware e/o software, una regione del fotogramma composito C comprendente dei pixel uguali a quelli dell'immagine sorgente. Nel caso in cui l'immagine composita è del tipo di Figura 1a o 1b, nella fase di copiatura è necessario dimezzare il numero dei pixel, o in orizzontale o in verticale. Le due immagini sinistra L e destra R inserite nel fotogramma composito C, pertanto, saranno deformate ed avranno una risoluzione orizzontale o verticale dimezzata.

In un fotogramma composito di un flusso video stereoscopico in un formato

alternativo come quello di Figura 1c, se le immagini ricevute in ingresso hanno dimensione 1280x720 pixel (cosiddetto formato 720p), che è uno dei formati utilizzati per la trasmissione di immagini a qualità migliorata, ma non propriamente ad alta definizione, allora un fotogramma composito adatto a contenerle entrambe è un fotogramma con 1920x1080 pixel, cioè un fotogramma di un flusso video di tipo 1080p (formato progressivo 1920 x 1080 pixel).

Nel caso in esame si ha il vantaggio che le due immagini destra R e sinistra L possono essere inserite nel fotogramma composito C senza essere deformate e/o sottocampionate, ma il prezzo che si paga è la necessità di suddividere una delle due immagini in almeno tre parti, come mostrato nell'esempio di Figura 1c.

Si noti che, nel formato alternativo di Figura 1c, la prima immagine può essere inserita in un punto qualsiasi del fotogramma contenitore (purché questo non ne richieda la scomposizione in aree); la scomposizione fotogramma della seconda immagine e l'inserimento nel fotogramma contenitore possono essere realizzati secondo una moltitudine di tecniche differenti, le cui caratteristiche non sono rilevanti ai fini della presente invenzione.

In ogni caso, il modo in cui le immagini sinistra L e destra R sono disposte all'interno del fotogramma contenitore C è influente ai fini dell'applicazione del metodo secondo l'invenzione.

In fase di codifica del flusso video stereoscopico, costituito dalla successione 101 dei fotogrammi contenitori C, viene inserito almeno un metadato M che contiene un'informazione relativa alla regione del fotogramma composito C che un decodificatore 2D che riceve in ingresso il flusso video stereoscopico 101 deve presentare in uscita dopo la decodifica.

I metadati M sono atti ad identificare una regione del fotogramma composito C

contenente un'immagine destinata all'occhio sinistro L oppure un'immagine destinata all'occhio destro R, purché detta immagine non sia stata scomposta in parti, come l'immagine destra dell'esempio di Figura 1c.

Se il flusso video stereoscopico 101 viene codificato secondo lo standard H.264/AVC, i metadati M possono vantaggiosamente sostituire i metadati relativi al "cropping rectangle", già previsto dallo standard stesso nella sezione 7.3.2.1.1 "Sequence parameter set data syntax" del documento ITU T H.264 del 03-2009 "Advanced video coding for generic audiovisual services".

Tali metadati erano stati originariamente introdotti in quanto la codifica H.264/AVC prevede la scomposizione dell'immagine in macroblocchi di 16x16 pixel, mentre 1080 non è un numero divisibile per 16. Pertanto il formato interno di rappresentazione dei singoli fotogrammi utilizzato da un codificatore può non coincidere con il formato effettivo dei fotogrammi (nel caso di fotogrammi 1920x1080 il codificatore H.264/AVC utilizza una rappresentazione su 1920x1088 pixel, aggiungendo quindi otto linee vuote). Tuttavia, i metadati che descrivono il "cropping rectangle" e previsti nello standard H.264/AVC, che servono a indicare al decodificatore quale porzione del fotogramma decodificato emettere in uscita per la riproduzione, permettono di definire in modo assolutamente generale una porzione rettangolare di fotogramma, che è quella da destinare alla visualizzazione.

In pratica, vengono definite le distanze dai bordi del fotogramma, espresse in numero di pixel, dei lati orizzontali e verticali della finestra di delimitazione che delimita la parte destinata alla visualizzazione.

In generale, è possibile definire la finestra di delimitazione dell'immagine in altri modi del tutto equivalenti; per esempio è possibile indicare la posizione di uno dei

vertici del rettangolo e le dimensioni dei lati, ovvero le coordinate di due vertici opposti.

Quindi, utilizzando in modo diverso i metadati M relativi al “cropping rectangle” già standardizzato ed usando quindi la stessa sintassi, è possibile specificare ad un decodificatore conforme alla norma H.264/AVC la riproduzione di un’area arbitraria del fotogramma composito C.

Con riferimento alla Figure 3a, 3b e 3c vengono indicate in modo tratteggiato alcune possibili regioni delimitate dalla finestra di delimitazione, la cui informazione relativa deve essere contenuta, secondo l’invenzione, in almeno un metadato M.

Ad esempio, nel fotogramma contenitore C di Figura 3a in formato side-by-side e nel fotogramma contenitore C di Figura 3c in formato alternativo, l’area coperta dalla finestra di delimitazione coincide con l’area occupata dall’immagine sinistra L del fotogramma contenitore C.

Come altro esempio, nel fotogramma contenitore C di Figura 3b in formato “top bottom”, l’area coperta dalla finestra di delimitazione coincide con l’area occupata dall’immagine destra R del fotogramma contenitore C.

Completata la fase di codifica, il flusso video stereoscopico può essere trasmesso su un canale di comunicazione e/o registrato su un opportuno supporto (ad es. CD, DVD, Blu-ray, memorie di massa, e così via).

In Figura 4a viene raffigurato un normale sistema per la fruizione di contenuti 2D, costituito da un decoder 1100 e da un display o dispositivo di visualizzazione 1110. Il decoder 1100 rende disponibili una delle due immagini destra R o sinistra L al dispositivo di visualizzazione 1110 (ad esempio un televisore), consentendo così la fruizione in 2D di contenuti 3D. Il decoder 1100 può essere o un set-top-box

separato dal televisore oppure può essere integrato nello stesso.

Considerazioni analoghe a quelle che vengono ora qui fatte si applicano ad un lettore (ad es. un lettore DVD) che legge un fotogramma contenitore e lo elabora al fine di ottenere un'immagine 2D.

Tornando ora alla Figura 4a, il sistema ricevente riceve (via cavo o antenna) un flusso video stereoscopico 1103 contenente fotogrammi composti C. Un decoder o decodificatore 1100 legge i metadati M che indicano quale porzione dei fotogrammi C deve essere utilizzata per la visualizzazione su un display o dispositivo di visualizzazione 1110. Successivamente, esso estrae l'immagine contenuta nella porzione di fotogramma composito C indicata dai metadati M e la invia al display 1110.

Può accadere che il formato dell'immagine in uscita dal decodificatore 1100 non sia tra quelli supportati in modo nativo dal display 1110. Per esempio, nel caso di un fotogramma 1080p in formato side-by-side, al momento della decodifica il ricevitore 1100 emetterà in uscita un fotogramma avente dimensione orizzontale pari alla metà della dimensione orizzontale del fotogramma ad alta definizione: l'area del fotogramma risultante sarà quindi pari a 960×1080 pixel ed il suo formato sarà 8/9.

Questo non è un formato immagine supportato in modo nativo dal display ma può essere riportato nel formato 1920×1080 per mezzo di opportune operazioni di interpolazione che possono essere effettuate da un dispositivo presente in tutti i display di tipo recente. Si tratta del dispositivo chiamato scaler, che ha la funzione di adattare il formato ricevuto a quello desiderato dall'utente (si pensi al fatto che tutti i televisori recenti sono dotati di un tasto del telecomando che permette di scegliere fra i formati 4/3, 16/9, 14/9, auto, ed al fatto che è possibile visualizzare

l'immagine in un formato diverso da quello nativo).

Nel caso in cui il fotogramma composito sia del tipo della Figura 1c, il formato dell'immagine compatibile 2D è del tipo 720p: si tratta di un formato standard che tutti i televisori HDTV o HD Ready sono in grado di gestire. Pertanto lo scaler sicuramente sarà in grado di espandere l'immagine compatibile affinché copra l'intero schermo.

Test effettuati su alcuni televisori hanno mostrato che anche nel caso in cui vengano usati i fotogrammi composti del tipo di Figura 1a oppure Figure 1b, lo scaler è in grado di visualizzare correttamente le immagini compatibili, riportandole cioè al formato 16/9 a tutto schermo.

La fruizione del video stereoscopico in modalità 3D sarà invece possibile solo utilizzando un decoder di nuova generazione che sia dotato di mezzi atti a riconoscere un'opportuna segnalazione contenuta nel flusso video stereoscopico e sia in grado di decodificare il video senza applicare la finestra di delimitazione.

Questo risultato può essere raggiunto in vari modi. Per esempio, si può aggiungere un ulteriore metadato N che specifica che la finestra di delimitazione è valida solo per i decoder/televisori 2D.

In tal modo, i prodotti esistenti (cioè 2D) non devono subire alcun aggiornamento o variazione di firmware/hardware.

In una forma di realizzazione preferita dell'invenzione, il metadato N che specifica se deve essere utilizzata la finestra di delimitazione originale o meno può essere semplicemente il metadato che specifica se la trasmissione è in 3D oppure in 2D.

La trasmissione di questo metadato N può ad esempio avvenire o in almeno una tabella SI (Service Information) oppure nei messaggi H.264 SEI (Supplemental Enhancement Information). In alternativa, si possono destinare una o più righe del

fotogramma composito alla trasmissione di metadati specifici per il 3D.

La Figura 4b illustra un sistema ricevente adatto alla visualizzazione di contenuti in 3D. Tale sistema si compone di tre parti, che possono costituire apparati indipendenti oppure possono essere accorpati in vario modo. Tali parti sono: un decoder 1100', un de-packer o spaccettatore 1105, ed un display 1110'.

Dal confronto con la figura 4a si evince che è necessario un componente aggiuntivo, e cioè lo spaccettatore 1105. Il decoder 1100' produce in uscita la sequenza di fotogrammi compositi C nella loro interezza, cioè ignorando le informazioni della finestra di delimitazione: si tratta quindi di un decoder non-standard, in quanto i decoder standard di arte nota dopo la decodifica del fotogramma, presentano in uscita solo la parte di immagine compresa nella finestra di delimitazione. Lo spaccettatore 1105 estrae dai fotogrammi compositi C le immagini L ed R e le presenta al display 1110' con le modalità previste dal display 1110' stesso.

Per fare ciò lo spaccettatore 1105 deve conoscere le modalità con cui è stato costruito il fotogramma composito. Uno spaccettatore 1105 adatto ad operare su vari formati e non su un formato unico deve pertanto leggere dei metadati P che specificano tali modalità.

Tali metadati P possono essere inseriti nelle tabelle SI o nei messaggi SEI oppure possono essere inseriti in una riga del fotogramma composito. Nel primo caso, è necessario che il decoder 1100' legga i suddetti metadati e li trasmetta allo spaccettatore 1105, ad esempio tramite un'interfaccia HDMI: si noti che la specifica di tale interfaccia è stata recentemente aggiornata proprio a questo scopo. Nel secondo caso la situazione è più semplice perché lo spaccettatore 1105 trova i metadati P direttamente nel fotogramma composito.

Dalla descrizione effettuata risultano pertanto chiare le caratteristiche della presente invenzione, così come chiari risultano i suoi vantaggi.

Un primo vantaggio della presente invenzione è quello per cui un broadcaster che voglia trasmettere un programma 3D, decodificabile anche da un decoder 2D, non deve trasmettere due flussi video simultanei, risparmiando in questo modo sulla banda necessaria per il trasporto del flusso video.

Un secondo vantaggio della presente invenzione è quello per cui il metodo è applicabile indipendentemente dal modo in cui i fotogrammi che trasportano le immagini destra e sinistra sono impacchettati all'interno di un fotogramma contenitore.

Numerose sono le varianti possibili al metodo per la generazione e la ricostruzione di un flusso video stereoscopico e relativi dispositivi di generazione e ricostruzione, senza per questo uscire dai principi di novità insiti nell'idea inventiva, così come è chiaro che nella sua attuazione pratica le forme dei dettagli illustrati potranno essere diverse, e gli stessi potranno essere sostituiti con degli elementi tecnicamente equivalenti.

Ad esempio, mentre nella descrizione si è fatto riferimento in particolare allo standard H.264/AVC, è chiaro che il metodo si applica anche quando la codifica è eseguita con altri standard, purché detti standard prevedano metadati simili a quelli usati per definire il cropping rectangle dello standard H.264/AVC.

Ad esempio, anche se nella descrizione si è fatto riferimento in particolare a fotogrammi composti di formato 1080p, è chiaro che l'invenzione prescinde dalle dimensioni del fotogramma composito e da quelle fotogrammi delle immagini componenti: l'invenzione si applica quindi anche al caso in cui il fotogramma composito è di dimensioni 720p, oppure è del tipo ad altissima risoluzione

(formati cosiddetti 2k e 4k, cioè con circa 2000 ovvero 4000 linee).

Dunque è facilmente comprensibile che la presente invenzione non è limitata ad un metodo per la generazione e la ricostruzione di un flusso video stereoscopico e relativi dispositivi, ma è passibile di varie modificazioni, perfezionamenti, sostituzioni di parti ed elementi equivalenti senza però allontanarsi dall'idea dell'invenzione, così come è precisato meglio nelle seguenti rivendicazioni.

RIVENDICAZIONI

1. Metodo per generare un flusso video stereoscopico digitale (101) comprendente fotogrammi contenitori (C), detti fotogrammi contenitori (C) comprendendo informazioni relative ad un'immagine destra (R) e ad un'immagine sinistra (L), caratterizzato dal fatto di inserire in fase di codifica di detto flusso video stereoscopico digitale (101) almeno un metadato (M) atto ad identificare una regione di un fotogramma contenitore (C) contenente una sola di dette due immagini (L,R).
2. Metodo secondo la rivendicazione 1, in cui il flusso video stereoscopico digitale è codificato secondo uno standard che prevede una finestra di delimitazione dell'immagine ed in cui detto almeno un metadato (M) definisce detta finestra di delimitazione e, pur utilizzando la stessa sintassi dello standard, si riferisce ad una sola delle due immagini (L,R) contenute in detto fotogramma contenitore (C).
3. Metodo secondo la rivendicazione 2, in cui detto standard è lo standard H.264/AVC.
4. Metodo secondo una delle rivendicazioni precedenti, in cui detta almeno una regione di detto fotogramma contenitore (C) presenta dimensioni tali da poter essere visualizzata senza operare alterazioni del rapporto larghezza/altezza di detta regione.
5. Metodo secondo una delle rivendicazioni da 1 a 3, in cui detta almeno una regione di detto fotogramma contenitore (C) viene interpolata e convertita di formato in modo tale da poter essere visualizzata a tutto schermo in formato 16/9.
6. Metodo secondo una delle rivendicazioni precedenti, in cui viene inserito nel flusso video stereoscopico digitale almeno un ulteriore metadato (N) che indica al decodificatore stereoscopico che deve ignorare detto almeno un metadato (M).

7. Metodo secondo la rivendicazione 6, in cui detto ulteriore metadato (N) viene trasmesso in almeno una tabella SI oppure nei messaggi SEI, relativi a detto flusso video stereoscopico digitale codificato secondo lo standard H.264/AVC.
8. Metodo secondo la rivendicazione 6 o 7, in cui detto almeno un ulteriore metadato (N) è il metadato che indica se il programma è in 2D oppure in 3D.
9. Dispositivo (100) per codificare un flusso video stereoscopico digitale (101) comprendente fotogrammi contenitori (C), detti fotogrammi contenitori (C) comprendendo informazioni relative ad un'immagine destra (R) e ad un'immagine sinistra (L), caratterizzato dal fatto di comprendere mezzi atti ad implementare il metodo secondo una qualsiasi delle rivendicazioni da 1 a 8.
10. Metodo per ricostruire almeno un'immagine compatibile 2D a partire da un flusso video stereoscopico digitale comprendente fotogrammi contenitori (C), detti fotogrammi contenitori (C) comprendendo informazioni relative ad un'immagine destra (R) e ad un'immagine sinistra (L), caratterizzato dal fatto di estrarre in fase di decodifica di detto flusso video stereoscopico digitale (101) almeno un metadato (M) atto ad identificare una regione di un fotogramma contenitore (C) contenente una sola di dette due immagini (L,R).
11. Metodo secondo la rivendicazione 10, in cui il flusso video stereoscopico digitale è decodificato secondo uno standard che prevede una finestra di delimitazione dell'immagine ed in cui detto almeno un metadato (M) definisce detta finestra di delimitazione e, pur utilizzando la stessa sintassi dello standard, si riferisce ad una sola delle due immagini (L,R) contenute in detto fotogramma contenitore (C).
12. Metodo secondo la rivendicazione 11, in cui detto standard è lo standard H.264/AVC.

13. Metodo secondo una delle rivendicazioni da 10 a 12, in cui detta almeno una regione di composita detto fotogramma contenitore (C) presenta dimensioni tali da poter essere visualizzata senza alterarne il rapporto larghezza/altezza.

14. Metodo secondo una delle rivendicazioni da 10 a 12, in cui detta almeno una regione di detto fotogramma contenitore (C) viene interpolata e convertita di formato in modo tale da poter essere visualizzata a tutto schermo in formato 16/9.

15. Decodificatore atto a ricostruire almeno un'immagine compatibile 2D a partire da un flusso video stereoscopico comprendente fotogrammi contenitori (C), detti fotogrammi contenitori (C) comprendendo informazioni relative ad un'immagine destra (R) e ad un'immagine sinistra (L), caratterizzato dal fatto di comprendere mezzi per implementare il metodo secondo una delle rivendicazioni da 10 a 14.

16. Metodo per decodificare un flusso video stereoscopico digitale ed estrarre le due immagini destra (R) e sinistra (L) di un fotogramma contenitore (C), caratterizzato dal fatto di leggere da parte di un decodificatore (1100') almeno un metadato (N) indicante se il flusso video è stereoscopico o meno e, in caso in cui detto flusso video sia stereoscopico:

- ignorare l'informazione contenuta in almeno un ulteriore metadato (M) che definisce una finestra di delimitazione;
- produrre la sequenza dei fotogrammi contenitori (C) nella loro integrità;
- inviare detta sequenza di fotogrammi contenitori (C) ad uno spaccettatore (1105) che produce dette due immagini destra (R) e sinistra (L) secondo le modalità richieste da un display stereoscopico (1110').

17. Metodo secondo la rivendicazione 16, in cui detto decodificatore (1100') invia a detto spaccettatore (1105) ulteriori metadati (P) indicanti le modalità di impacchettamento delle immagini sinistra (L) e destra (R), in particolare contenuti

nel campo SI e/o nel messaggio SEI relativi a detto flusso video stereoscopico.

18. Metodo secondo la rivendicazione 16, in cui detto spaccettatore (1105) legge ulteriori metadati (P) contenuti in un fotogramma contenitore ed indicanti le modalità di impacchettamento delle immagini sinistra (L) e destra (R).

19. Metodo secondo la rivendicazione 17 o 18, in cui detti ulteriori metadati (P) contengono finestre di delimitazione relative a ciascuna regione del fotogramma composito (C).

20. Metodo secondo la rivendicazione 16, in cui se detto metadato (N) indica che detto flusso video non è stereoscopico, il decodificatore (1100') legge l'informazione contenuta in detto almeno un ulteriore metadato (M) che definisce detta finestra di delimitazione, ed invia al display (1110) l'immagine definita da detta finestra di delimitazione.

21. Decodificatore atto a decodificare un flusso video stereoscopico digitale ed estrarre le due immagini destra (R) e sinistra (L) di un fotogramma contenitore (C), comprendente mezzi per implementare il metodo secondo una o più delle rivendicazioni da 16 a 19.

22. Decodificatore atto a decodificare un flusso video stereoscopico digitale ed estrarre un'immagine di un fotogramma contenitore (C), comprendente mezzi per implementare il metodo secondo la rivendicazione 20.

23. Spaccettatore (1105) atto a produrre le immagini destra e sinistra di un flusso video stereoscopico secondo le modalità richieste da un display stereoscopico secondo una o più delle rivendicazioni da 16 a 19.

24. Flusso video stereoscopico (1101) caratterizzato dal fatto di comprendere almeno un fotogramma contenitore (C) ed almeno un metadato (M,N) secondo una qualsiasi delle rivendicazioni da 1 a 8.

CLAIMS

1. Method for generating a digital stereoscopic video stream (101) comprising container frames (C), said container frames (C) comprising information about a right image (R) and a left image (L), characterized in that, when coding said digital stereoscopic video stream (101), at least one metadatum (M) is entered which is adapted to identify a region of a container frame (C) containing only one of said two images (L,R).
2. Method according to claim 1, wherein the digital stereoscopic video stream is coded according to a standard which includes an image delimiting window, and wherein said at least one metadatum (M) defines said delimiting window and, while using the same syntax as the standard, refers to only one of the two images (L,R) contained in said container frame (C).
3. Method according to claim 2, wherein said standard is the H.264/AVC standard.
4. Method according to any of the preceding claims, wherein the dimensions of said at least one region of said container frame (C) are such that it can be displayed without requiring any changes to the width/height ratio of said region.
5. Method according to any of claims 1 to 3, wherein said at least one region of said container frame (C) is interpolated and format-converted so that it can be displayed in full screen in the 16/9 format.
6. Method according to any of the preceding claims, wherein at least one additional metadatum (N) is entered into the digital stereoscopic video stream to indicate to the stereoscopic decoder that said at least one metadatum (M) must be ignored.
7. Method according to claim 6, wherein said additional metadatum (N) is sent in

at least one SI table or in the SEI messages concerning said digital stereoscopic video stream coded in accordance with the H.264/ AVC standard.

8. Method according to claim 6 or 7, wherein said at least one additional metadatum (N) is the metadatum which indicates whether the program is in 2D or 3D.

9. Device (100) for coding a digital stereoscopic video stream (101) comprising container frames (C), said container frames (C) comprising information about a right image (R) and a left image (L), characterized by comprising means adapted to implement the method according to any of claims 1 to 8.

10. Method for rebuilding at least one 2D-compatible image starting from a digital stereoscopic video stream comprising container frames (C), said container frames (C) comprising information about a right image (R) and a left image (L), characterized in that, when decoding said digital stereoscopic video stream (101), at least one metadatum (M) is extracted which is adapted to identify a region of a container frame (C) containing only one of said two images (L,R).

11. Method according to claim 10, wherein the digital stereoscopic video stream is decoded according to a standard which includes an image delimiting window, and wherein said at least one metadatum (M) defines said delimiting window and, while using the same syntax as the standard, refers to only one of the two images (L,R) contained in said container frame (C).

12. Method according to claim 11, wherein said standard is the H.264/ AVC standard.

13. Method according to any of claims 10 to 12, wherein the dimensions of said at least one region of said container frame (C) are such that it can be displayed without requiring any changes to the width/height ratio thereof.

14. Method according to any of claims 10 to 12, wherein said at least one region of said container frame (C) is interpolated and format-converted so that it can be displayed in full screen in the 16/9 format.

15. Decoder adapted to rebuild at least one 2D-compatible image starting from a stereoscopic video stream comprising container frames (C), said container frames (C) comprising information about a right image (R) and a left image (L), characterized by comprising means for implementing the method according to any of claims 10 to 14.

16. Method for decoding a digital stereoscopic video stream and extracting the two right (R) and left (L) images of a container frame (C), characterized in that a decoder (1100') reads at least one metadatum (N) indicating whether the video stream is stereoscopic or not and, if said video stream is stereoscopic:

- the information contained in at least one additional metadatum (M) defining a delimiting window is ignored;
- the sequence of container frames (C) is produced integrally;
- said sequence of container frames (C) is sent to a de-packer (1105), which produces said two right (R) and left (L) images as required by a stereoscopic display (1110').

17. Method according to claim 16, wherein said decoder (1100') sends to said de-packer (1105) additional metadata (P) indicating how the left (L) and right (R) images have been packed, in particular contents of the SI field and/or of the SEI message pertaining to said stereoscopic video stream.

18. Method according to claim 16, wherein said de-packer (1105) reads additional metadata (P) contained in a container frame and indicating how the left (L) and right (R) images have been packed.

19. Method according to claim 17 or 18, wherein said additional metadata (P) contain delimiting windows for each region of the composite frame (C).
20. Method according to claim 16, wherein if said metadatum (N) indicates that said video stream is not stereoscopic, then the decoder (1100') reads the information contained in said at least one additional metadatum (M) that defines said delimiting window and sends the image defined by said delimiting window to the display (1110).
21. Decoder adapted to decode a digital stereoscopic video stream and extract the two right (R) and left (L) images of a container frame (C), comprising means for implementing the method according to one or more of claims 16 to 19.
22. Decoder adapted to decode a digital stereoscopic video stream and extract an image of a container frame (C), comprising means for implementing the method according to claim 20.
23. De-packer (1105) adapted to produce the right and left images of a stereoscopic video stream as required by a stereoscopic display according to one or more of claims 16 to 19.
24. Stereoscopic video stream (1101) characterized by comprising at least one container frame (C) and at least one metadatum (M,N) according to any of claims 1 to 8.

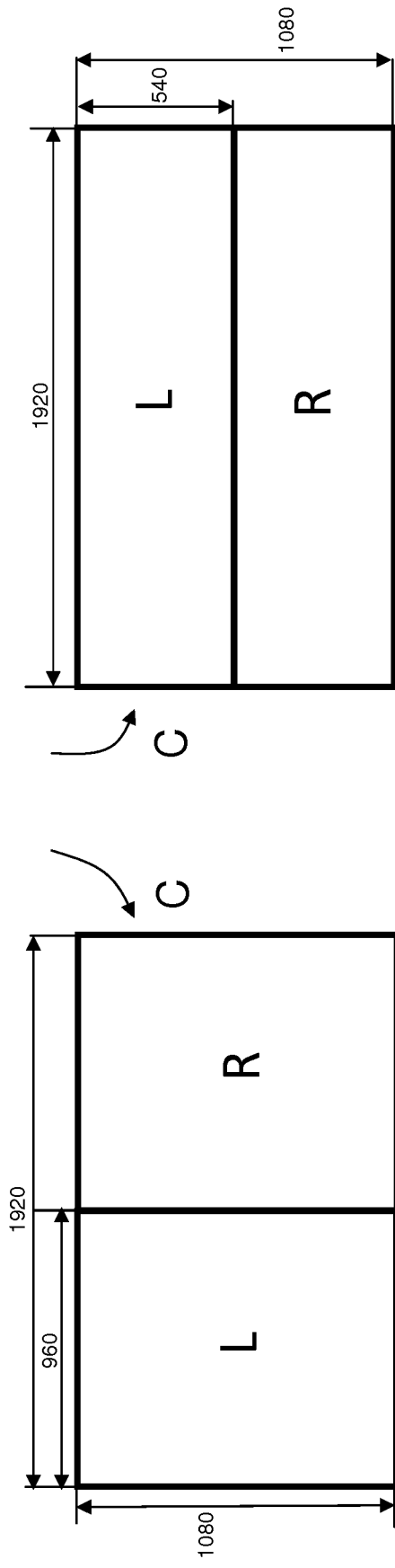


Fig. 1a

Fig. 1b

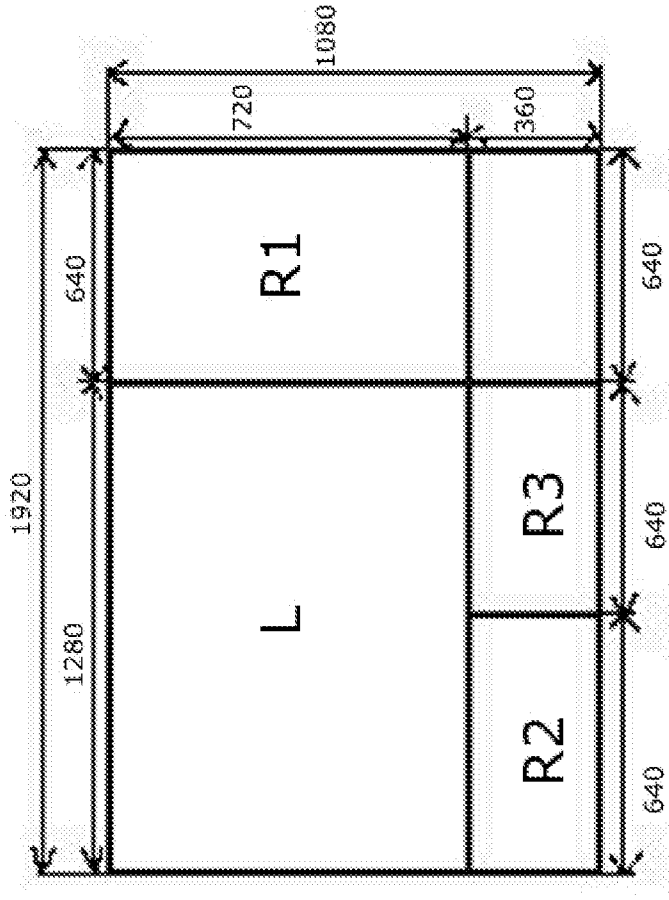


Fig. 1c

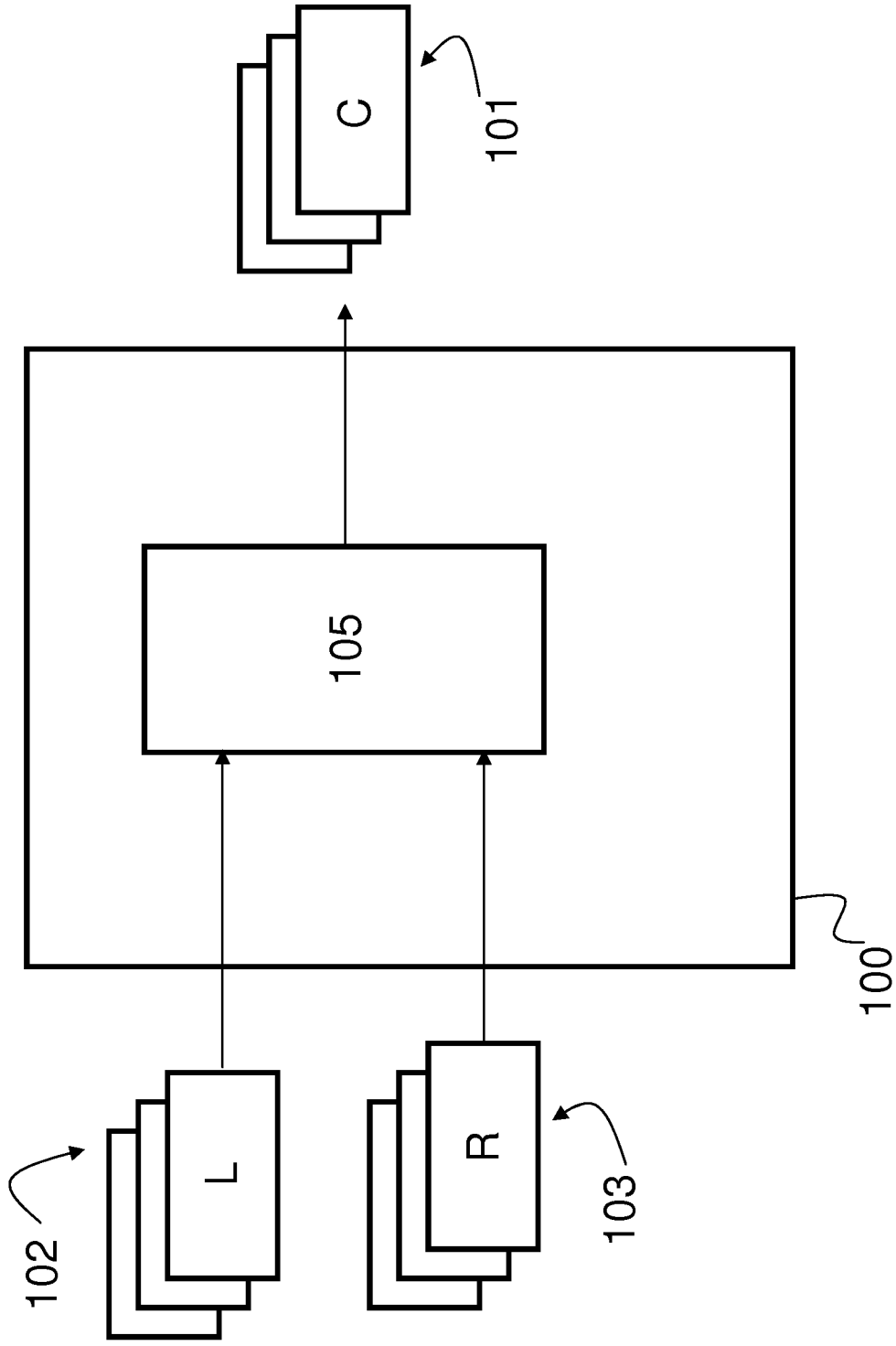


Fig. 2

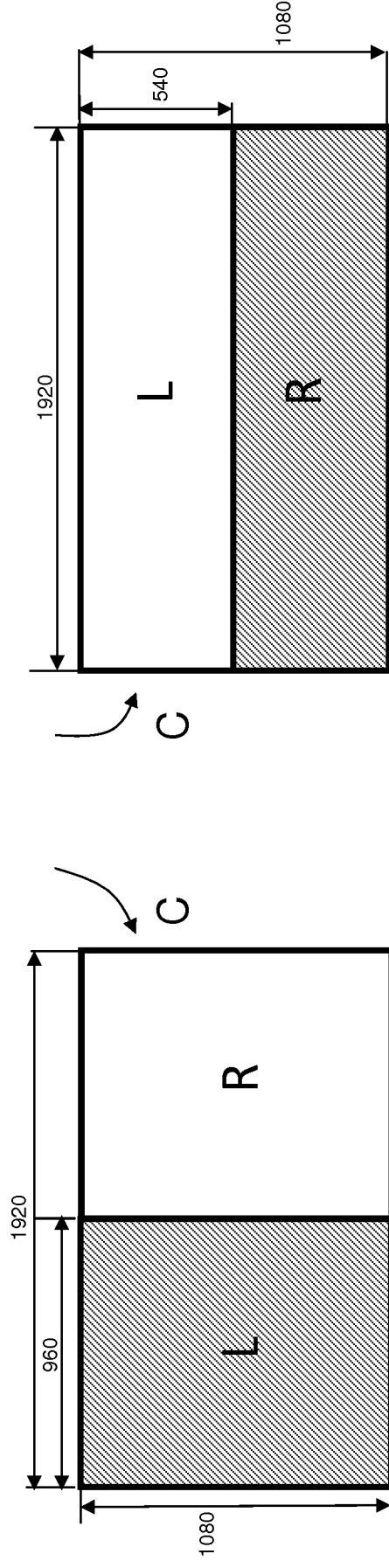


Fig. 3a

Fig. 3b

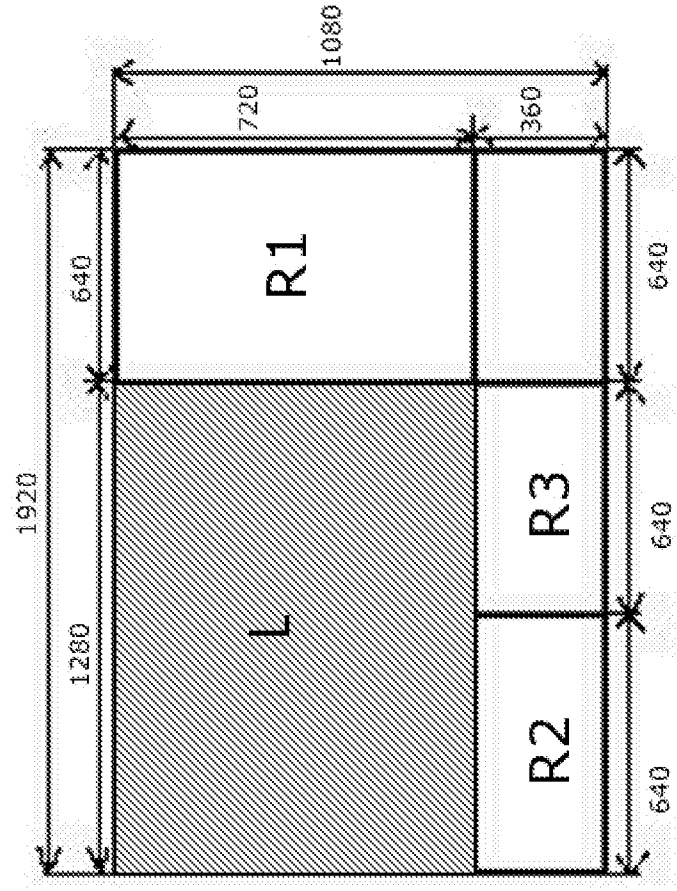


Fig. 3c

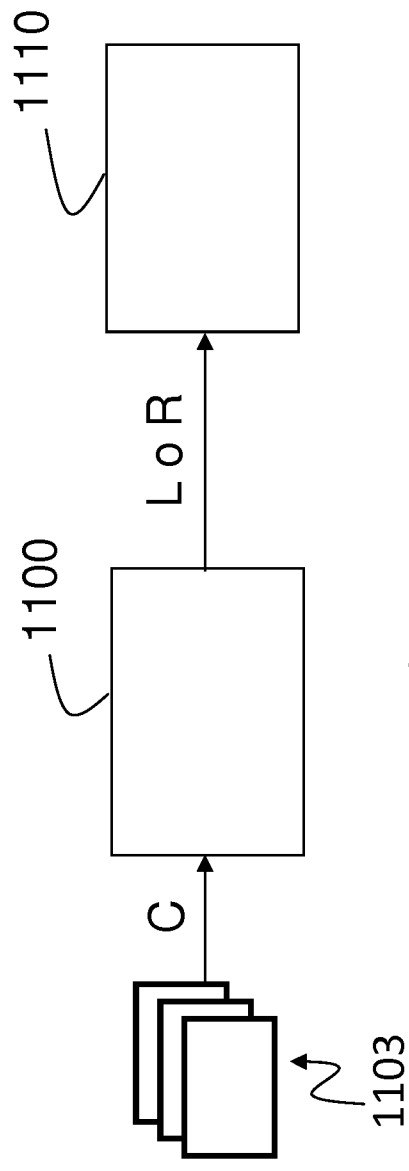


Fig. 4a

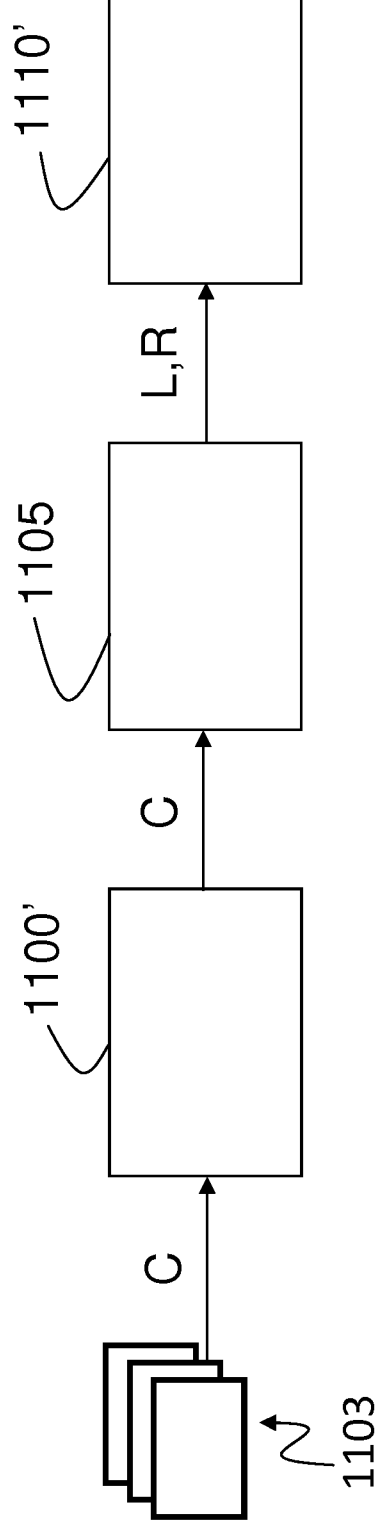


Fig. 4b